

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 2, 2004, 18:15:03 ; Search time 9476 Seconds

(without alignments)

5228.059 Million cell updates/sec

Title: US-09-720-934-2

Perfect score: 5895

Sequence: 1 MAQFPTFGSLDIWAIVE.....QVGLFPSYVXKLTMDPDSQ 1143

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp

-Q=/cgn2_1/USPto_spool_p/US09720934/runat_29072004.164338.1218/app.query.fasta_1.1287

-DB=GenEmbl -QFMT=fastap -SUFFIX=jul29.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=numan40.cdi -LIST=45

-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09720934 @CGN 1.1 6972 @runat_29072004.164338.1218 -NCPU=6 -ICPU=3

-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.ov.*

5: gb.pat.*

6: gb.ph.*

7: gb.pl.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sy.*

12: gb.un.*

13: gb.vi.*

14: em.ba.*

15: em.fun.*

16: em.hum.*

17: em.in.*

18: em.mu.*

19: em.om.*

20: em.or.*

21: em.pat.*

22: em.ph.*

23: em.pl.*

24: em.ro.*

25: em.sts.*

26: em.un.*

27: em.vi.*

28: em.yv.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	5895	100.0	5199	6	BD205033	Isolated
2	5849.5	99.2	5195	6	BD205035	Isolated
3	5837	99.0	5381	9	AF114488	Homo sapi
4	5837	99.0	5458	6	BD205034	Isolated
5	5837	99.0	6439	9	AF114487	Homo sapi
6	5829	98.9	5287	9	AF064243	Homo sapi
7	5829	98.9	7247	9	AF064244	Homo sapi
8	5460	92.6	3812	10	AF132672	Rattus no
9	5437.5	92.2	3723	10	AF132478	Mus muscu
10	5437.5	92.2	5145	10	AF132481	Mus muscu
11	5414.5	91.8	4025	10	AF127798	Rattus no
12	5057	85.8	4321	9	BSM06384	Homo sapi
13	4754	80.6	4103	5	AF032118	Xenopus l
14	3344	56.7	3231	6	BD205037	Isolated
15	3246	55.1	2079	6	BD205036	Isolated
16	3209	54.4	2131	6	BD127168	Primer fo
17	3209	54.4	2131	9	AK074554	Homo sapi
18	3183	54.0	2199	9	BC058925	Homo sapi
19	2862.5	50.3	2131	6	AX880753	Sequence
20	2862.5	50.3	2131	6	BD158570	Primer fo
21	2862.5	50.3	2131	9	AK027846	Homo sapi
22	2837	48.1	5828	6	BD167848	Method fo
23	2837	48.1	5938	9	AF248540	Homo sapi
24	2837	48.1	5938	9	AB033082	Homo sapi
25	2813.5	47.7	4557	9	AF182199	Homo sapi
26	2813.5	47.7	6091	9	AF182198	Homo sapi
27	2779	47.1	4977	10	AF132479	Mus muscu
28	2776	47.1	3241	9	HSU61166	Human SH3 d
29	2772	47.0	3594	10	AF132480	Mus muscu
30	2740.5	46.5	4447	6	AX428899	Sequence
31	2388	40.5	1676	6	BD127640	Primer fo
32	2388	40.5	1676	9	AK075290	Homo sapi
33	2371	40.2	5804	10	AK122480	Mus muscu
34	1930.5	32.7	113146	2	AC138440	Tetradon
35	1756	29.8	1299	9	BC039036	Homo sapi
36	1728.5	29.3	2017	6	BD264687	Compositi
37	1728.5	29.3	2017	6	AR238090	Sequence
38	1728.5	29.3	2017	6	AR257631	Sequence
39	1728.5	29.3	2017	6	AR283677	Sequence
40	1728.5	29.3	2017	6	AX366305	Sequence
41	1691.5	28.7	3750	3	AF054612	Drosophil
42	1669	28.3	3999	3	AF053957	Drosophil
43	1616.5	27.4	136937	2	AC139627	Takifugu
44	1564	26.5	2085	9	BC038963	Homo sapi
45	1558	26.4	1749	9	BC020921	Homo sapi

ALIGNMENTS

RESULT 1

```

BD205033      5199 bp      DNA      linear      PAT 17-JUL-2003
LOCUS          Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION     leukemia and utilization thereof.
ACCESSION      BD205033
VERSION        JP 2002511267-A/1.
KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        1 (bases 1 to 5199)
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          Korenberg, J.R. and Chen, X.N.
                Isolated SH3 gene relating to myeloproliferative disorders and
                leukemia and utilization thereof
JOURNAL        Patent: JP 2002511267-A 1 16-APR-2002;
                CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT        OS Homo sapiens (human)
                PN JP 2002511267-A/1
                PD 16-APR-2002
                PF 16-APR-1999 JP 2000543610
                PR 16-APR-1998 US 60/082007
                PI JULIE R KORENBERG, XIAO NING CHEN
                PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
                PC C12Q1/68,
                PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
                CC Isolated SH3 gene relating to myeloproliferative disorders and
                CC leukemia
                CC and utilization thereof.
                FH Key      Location/Qualifiers
                FT source    1. .5199
                FT          /organism='Homo sapiens (human)'.
                FT          Location/Qualifiers
                FT          1. .5199
                FT          /organism='Homo sapiens'
                FT          /mol_type='genomic DNA'
                FT          /db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.:      2,19e-231      Length:      5199
Score:          5895.00      Matches:    1143
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6      Gaps:      0

US-09-720-934-2 (1-1143) x BD205033 (1-5199)

Qy      1      MetAlaGlnPheProThrPheGlyGlySerLeuAspIleTrrAlaIleThrValGlu 20
Db      208      ATGGCTCGTTTCCACACCTTTTGGTGGAGCCCTGGATATCTGGGCCATTAACGTAGAG 267
Qy      21      GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db      268      GAAAGAGCGAAGCATGATCAGCAGTCCATAGTTTAAAGCCAATATCTGGATTCAATTACT 327
Qy      41      GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db      328      GGTGATCAAGCTAGAACTTTTTCATCTGGGTTTACCTCAACCTGTTTATGACAG 387
Qy      61      IleTrrAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db      388      ATATGGGCACCTAGCTGACATGAATATGATGGAGATGGATGATCAAGTGGAGTTTCCATA 447
Qy      81      AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db      448      GCTATGAACTTATCAACTGAAGCTACAAGGATATCAGCTACCTCTGCACTTCCCCCT 507
Qy      101      ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyIle 120
Db      508      GTCATGAAACAGCAACCAAGTGTGCTATTCTAGCCAGCCAGCATTTGGTATGGAGGTATC 567

```

```

Qy      121      AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db      568      GCCAGCATGCCACCGCTTACAGCTGTCTCCAGTGCATATGGATGCCATTCACATGTGT 627
Qy      141      GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db      628      GGAATGCTCCAACCCCTAGTATCTCTGTCCACAGCAGCTGTGCCCCCTGGCTAAC 687
Qy      161      GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db      688      GGGGCTCCCTGTTTATACAACTCTGCTCATTTGCTCATCTCTGACGACACATGGCA 747
Qy      181      LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db      748      AAGAGTTCTTCTCTTAGTAGATCTGCTCCAGGTCACAACTAAACACTAAATTTACA 807
Qy      201      AlaGlnSerPheAspValAlaSerValProProValAlaGluTrrAlaValProGlnSer 220
Db      808      GCACAGTCATTTGATGTGGCCAGTGTCCACAGTGGCAGAGTGGCTGTCTCAGTCA 867
Qy      221      SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db      868      TCAAGACTGAAATACAGCAATTTATTCATATAGTCAAACTATGAGTGACACTTA 927
Qy      241      ThrGlyProGlnAlaAtgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
Db      928      ACAGTCCCCAAGCAAGAACTATTCTTATGAGTCAAGTTTACCACAGCTCAGCTGGCT 987
Qy      261      SerIleTrrAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
Db      988      TCAATATGGAATCTTCTGACATTCATCAAGATGGAAACTTACAGCAGAGAGAAATTC 1047
Qy      281      LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
Db      1048      CTGGCAATGCACTCAITGATGTAGTATGTCTGGCCAACTGCTGCACTGTCTGCTGCT 1107
Qy      301      ProGluTrrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
Db      1108      CCAGATACATTCACCTCTTTTAGAAGAGTTGATCTGCGAGTGGTATATCTATCTCAT 1167
Qy      321      SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
Db      1168      AGCTCAACATCTGTAGATCAGAGGCTACACAGAGGAACAGATTTTAGAAGATGAACA 1227
Qy      341      GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
Db      1228      CAATTAGAAAAGAAATTAACCTGTAACTGTTTGAAGATAAGACCGCGAGAACTTTGA 1287
Qy      361      GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGlnGlnArgLysGluGln 380
Db      1288      GGCAACCTGGAACTGGAGAAACGAAGCAAGCTCTCTCTGGAAACAGCAGCGCAGAG 1347
Qy      381      GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu 400
Db      1348      GAGCGCTTGGCCAGCTGGAGCGGCGAGCAGGAGAGAAAGAGCGTGGAGCGCGAGAG 1407
Qy      401      GlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
Db      1408      CAAGAGCGCAAAAGACAACTGGAATCTGGAAGCAACTTGGAAAGACGCGGAGCTAGA 1467
Qy      421      ArgGlnArgGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu 440
Db      1468      CGGCAGACAGAGAGAGAGAGAGAGAGAAATTCAGAGCGCGAGAGGCTGCAAAACGG 1527
Qy      441      LeuGluArgGlnArgGlnLeuGluTrrPgluArgAsnArgArgGlnGlnLeuLeuAsnGln 460
Db      1528      CTTGAAAGGCAACGACACTTTGAGTGGCAACGGAATCGAAGCGCAAGACTACTAAT 1587
Qy      461      ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysThrLeuGluPhe 480
Db      1588      AGAAACAAAGACAGAGGACATAGTTGTTACTGAAAGCAAGAAAGAAAGACTTTGAA 1647

```

QY	481	GluLeuGluAlaLeuAsnAspLysHisGlnLeuGluGlyLysLeuGlnAspIleArg	500
Db	1648	GAATTTAGAGGCTCTAAATGATATAAAGCATCACTAGAGGGAACCTTCAAGATATCAGA	1707
QY	501	CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg	520
Db	1708	TGTGCAITTGACCAACCAAGGCAAGAAATTGAGACACAAACAAATCTAGAGATTGAGA	1767
QY	521	IleAlaGluIleThrHisLeuGlnGlnLeuGlnGlnGluSerGlnGlnMetLeuGlyArg	540
Db	1768	ATTGCCGAATACCCATCTACAGCAACAAATACAGGAATCTCAGCAATGCTTGGAGA	1827
QY	541	LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu	560
Db	1828	CTTATTTCCAGAAAAACAGATACTCAATGACCACAAATTAACAAAGTTTCCAGCAGAACAGTTTG	1887
QY	561	HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln	580
Db	1888	CACAGAGATTCTACTTGTACCTTAAAGAGGCTTAGAAGCAAAAGAACTAGCTGGCAG	1947
QY	581	HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle	600
Db	1948	CACCTACGAGCACTGATGATGAGTGAGAGAAAGAACTAGATCAAAACTACAGGAGATT	2007
QY	601	AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln	620
Db	2008	GATAATTTTCAATTAATCAGCTGAAGGAACCTAAGAGAAATACAAATTAAGCAACAACTCCAG	2067
QY	621	LysGlnLysSerMetClnAlaGluArgLeuLysGlnLysGluClnGluArgLysIleIle	640
Db	2068	AAGCAAAAGTCCATGGAGGCTGAACACTGAACAGAAAGAAACAGAAAGAAATCATA	2127
QY	641	GluLeuGluLysGlnLysGluAlaGlnArgAlaGlnGluArgAspLysGlnTrp	660
Db	2128	GAATTAGAAAAACAAAAAGAGNAGCCCAAGACGAGCTCAGAAAGGACCAAGCAGTGG	2187
QY	661	LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGlu	680
Db	2188	CTGGAGCATGTGCAGCAGGAGGACGAGCATCAGAGACCAAGAAAACTCCACGAAGAGGAA	2247
QY	681	LysLeuLysArgGluGluSerValLysLysLysAspGlyGluClnLysGlyLysGlnGlu	700
Db	2248	AAACTGAAAGGGAGAGAGTGTCAAAAAAGAGGATGGCGAGAAAGAAAGCAACAGGAA	2307
QY	701	AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal	720
Db	2308	GCACAGACAAGCTGGTGGCTTTTCCATCAACACCAAGAACCAAGCTAAGCCAGCTGTC	2367
QY	721	GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal	740
Db	2368	CAGGCACCTGGTCCACTGCAGAAAAAGGTCCACTTACCATTCTGCACAGGAAAAATGTA	2427
QY	741	LysValValTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr	760
Db	2428	AAAGTGGTGATTACCGGGACCTGTACCCCTTTGAATCCAGAGCCCATGATGAATCACT	2487
QY	761	IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeu	780
Db	2488	ATCCAGCCAGGACATAGTGTGATGGTGAAGAAAGCCAACTGGAGAACCCCGCTGCTT	2547
QY	781	GlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIlePro	800
Db	2548	GGAGGAGAAATTAAGAGAAAGACAGGGTGGTTCCCTGCCAAACTATGCAGAGAAAAATCCCA	2607
QY	801	GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro	820
Db	2608	GAAAAATGAGGTTCCTCCAGTGAAACCCAGTGAATCAACATCTGCCCTTCCCTCC	2667
QY	821	LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr	840
Db	2668	AAACTGGCTTGGTGAGACCCCGCCCTTTGGCAGTAACTCTTCAGAGCCCTCCACG	2727
QY	841	ThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysPro	860

Db	2728	ACCCCTAATAACTGGCGGACTTCAGCTCCAGTGGCCACCAGCAGCAATGAGAAACCA	2787
QY	861	GluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThrValProSerAlaGly	880
Db	2788	GAACCGGATACTGGATGTCATGGCAGCCAGCCCTCTCTCACCGTTTCCAAGTCCCGC	2847
QY	881	GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro	900
Db	2848	CAGTTAAGGCAGAGGTCGCGCTTTACTCCAGCAAGCCACTGGCTCTCCCGCTCTCT	2907
QY	901	ValLeuGlyGlnGlyGluLysValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrpArg	920
Db	2908	GTGCTAGGCCAGGGTGAAGGTGGAGGGCTTACAGCTCAAGCCCTATATCTCTGGAGA	2967
QY	921	AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGln	940
Db	2968	GCCAAAAAGACCAACCACTTAATTTTAAACAAAAATGATGTCATCACCGTCCCTGGAACAG	3027
QY	941	GlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrpPheProLysSerTyr	960
Db	3028	CAAGACATGTGTGGTTTGGAGAAAGTTCAAGGTCAGAGGGTTGGTTCCCAAGTCTTAC	3087
QY	961	ValLysLeuIleSerClyProIleArgLysSerThrSerMetAspSerGlySerSerGlu	980
Db	3088	GTGAACACTCATTTTACGGGCCCATAGGAAGTCTTCAAGCATGGATTCTGGTCTTCAGAG	3147
QY	981	SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu	1000
Db	3148	AGTCTGTCTAGTCTAAAGCGAGTAGCTCTCCAGCAGCAAGCGGTCTGTTTCGGGAGAA	3207
QY	1001	GluIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAla	1020
Db	3208	GAATTTGCCAGGTATTTCCTCATACCCGCCACCGCGCCCGCAGCAGCTCACTCTCGCC	3267
QY	1021	ProGlyGlnLeuIleLeuIleArgLysLysAsnProGlyGlyTrpTrpGluGlyGluLeu	1040
Db	3268	CCTGGTCAGCTGATTTTGATTCGAAAAAGAACCCAGGTGGATGGTGGGAAGAGAGCTG	3327
QY	1041	GlnAlaArgGlyLysLysArgGlnIleGlyTrpPheProAlaAsnTyrValLysLeuLeu	1060
Db	3328	CAAGCACGTGGGAAAAAGGCCAGATAGCTGGTTCCTCCAGCTAATATATGTAAGCTTCTA	3387
QY	1061	SerProGlyThrSerLysIleThrProThrGluProProLysSerThrAlaLeuAlaAla	1080
Db	3388	AGCCCTGGGACGAGCAAAATCACTCCACAGAGCCACCTAAGTCAACAGCATTAGCGGCA	3447
QY	1081	ValLysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAlaPhe	1100
Db	3448	GTGTGCCAGGTGATGGGATGTACGACTACCCGCGCAATGACATGAGCTGGCCTTC	3507
QY	1101	AsnLysGlyGlnIleAsnValLeuAsnLysGluAspProAspTrpTrpLysGlyGlu	1120
Db	3508	AACAAGGCCAGCATCATCAACGTCCTCAACAGAGGACCTGACTGGTGGAAAGGAGAA	3567
QY	1121	ValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAsp	1140
Db	3568	GTCAATGGACAGTGGGGCTCTTCCCATCCAATTATGTGAAGCTGACCACAGCATGGAC	3627
QY	1141	ProSerGln	1143
Db	3628	CCAAGCCAG	3636

RESULT 2

LOCUS	BD205035	5195 bp	DNA	linear	PAT 17-JUL-2003			
DEFINITION	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.							
ACCESSION	BD205035							
VERSION	BD205035.1	GI:33014805						
KEYWORDS	JP 2002511267-A/3.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 5195)
 Korenberg, J.R. and Chen, X.N.
 Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 Patent: JP 2002511267-A 3 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL
 OS Homo sapiens (human)
 PN JP 2002511267-A/3
 PD 16-APR-2002
 PR 16-APR-1999 JP 2000543610
 PI JULIE R. KORENBERG, XIAO NING CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10, C12Q1/68,
 PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and leukemia
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..5195
 FT /organism='Homo sapiens (human)'.
 FEATURES
 source Location/Qualifiers
 1..5195
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.56e-229 Length: 5195
 Score: 5849.50 Matches: 1143
 Percent Similarity: 94.15% Conservative: 0
 Best Local Similarity: 94.15% Mismatches: 0
 Query Match: 99.23% Indels: 71
 DB: 6 Gaps: 1
 US-09-720-934-2 (1-1143) x BD205035 (1-5195)
 Qy 1 MetAlaGlnPheProThrPheGlySerLeuAspIleThrValGlu 20
 Db 239 ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCTGGGATATCTGGGCCATAAATCTAGAG 298
 Qy 21 GluArgAlaIysHisAspGlnGlnPheHisSerLeuIysProIleSerGlyPheIleThr 40
 Db 299 GAAAGAGCGAAGCATGATCAGCATGTTTCAATAGTTTAAAGCCAAATATCTGGATTCTATTACT 358
 Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 359 GGTGATCAAGCTAGAACTTTTTCATCTGGTTTACCTCAACCTGTTTACGACAG 418
 Qy 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 419 ATATGGGCACCTAGCTGACATGAATTAATGATGAAGATGGATCAAGTGGAGTTTCCATA 478
 Qy 81 AlaMetIysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 479 GCTATGAACACTTATCAAACTGAAGCTAAGGATATCAGCTTACCTCTGCACTTCCCTCC 538
 Qy 101 ValMetIysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 Db 539 GTCATGAACACAGCAACCACTGCTATTTCTAGCCGCCAGCATTTGGTATGGGAGGTATC 598
 Qy 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 599 GCCAGCATGCCACCGCTTACAGCTGTTGCTCCAGTGGCAATGGGATCCATTCACAGTTGTT 658
 Qy 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 Db 659 GGATGTCTCCACCTAGTATCTTCTGTTCCACAGCAGCTGTGCCCCCTCGGCTAAC 718

Qy 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 Db 719 GGGGCTCCCTCTGTTATACAACTCTCCCTGCACTTCTGCTATCTCTGAGCCACATGGCA 778
 Qy 181 LysSerSerPheSerArgSerGlyProGlySerGlnLeuAenThrLysLeuGlnIys 200
 Db 779 AAGAGTTCTTCTTTAGTAGATCTGGTCCAGGGTCCAACTAAACACTAAATTAACAAAG 838
 Qy 201 AlaGlnSerPheAspValAlaSerValProProValalagluTrpAlaValProGlnSer 220
 Db 839 GCACAGTCATTTGATGTGGCCAGTGTCCACCACTGGCAGAGTGGGTGTTCTCCAGTCA 898
 Qy 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 Db 899 TCAAGACTGAATACAGGCAATTTATTCATAGTCTATGACAAACATATGATGGACACTTA 958
 Qy 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
 Db 959 ACAGGTCCCAAGCAAGAACTATTCTTATGCAGTCAAGTTTACCACAGGCTCAGCTGGCT 1018
 Qy 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
 Db 1019 TCAATATGGAATCTTCTGACATTTGATCAAGATGGAACACTTACAGCAGAGAAATTTATC 1078
 Qy 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
 Db 1079 CTGGCAATGCACCTCATTTGATGTAGTATGTCTGCCCAACCACTGCCACCTGTCTGCT 1138
 Qy 301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
 Db 1139 CCAGATACATTCACCTTCTTTTAGAAGAGTTCCGATCTGGCAGTGTATATCTGTCTATA 1198
 Qy 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuLeuAspGluGlnGln 340
 Db 1199 AGCTCAACATCTGTAGATCAGAGCTACCAGAGGAACACAGTTTGTAGAGATGAACAA 1258
 Qy 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
 Db 1259 CAATTGACAAAGAAATTTACCTGTAACGTTTGAAGATAAGAGCGGAGAACTTTGAACGT 1318
 Qy 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGlnGlnGlnArgLysGluGln 380
 Db 1319 GGCAACCTGGAACTGGAGAAACGAAGCAAGCTCTCTCTGGAACAGCAGCGCAGGAGCAG 1378
 Qy 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGluGlu 400
 Db 1379 GAGCGCTGGCCCTGAGCTGGAGCGGCGGAGCAGGAGAGAGAGAGCTGAGCGCCAGAG 1438
 Qy 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
 Db 1439 CAAGAGCGCAAAAGACAACTGGAACTGGAGAAACCACTGGAAAAGCAGCGGAGCTAGAA 1498
 Qy 421 ArgGlnArgGluGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu 440
 Db 1499 CGGCAGAGAGAGAGAGAGAGAGAGAAATTCAGAGCGGAGAGGCTGCAAAACGGGAA 1558
 Qy 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGluLeuLeuAsnGln 460
 Db 1559 CTTGAAAGGCACACCAACTTTGAGTGGACGGATCGAAGGCAAGAACTACTTAATCA 1618
 Qy 461 ArgAsnLysGlnGlnGluAspIleValLeuLysAlaLysLysLysThrLeuGluPhe 480
 Db 1619 AGAAACAAAGAAACAAAGAGGACATAGTTGCTACTGAAAGCAAGAAAGAAAGACTTTTGAATTT 1678
 Qy 481 GluLeuGluAlaLeuAsnAspLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
 Db 1679 GAATTAGAGCTCTTAATGATAAAGCAATCACTAGAGGAACTTCAAGATATCAGA 1738
 Qy 501 CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg 520
 Db 1739 TGTGATTTGACCCCAAGCAAGAAATTTAGAGACCAACAACTCTAGAGAGTTGAGA 1798
 Qy 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnSerGlnMetLeuGlyArg 540

Db 1799 ATTGCGGAATCAATCCATCTACAGCAACAAATTAACAGAAATCTCAGCAAAATGCTTGGAGA 1858
QY 541 LeuileProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1859 CTTATTTCCAGAAAACAGATCTCAATGACCAATTAATAACAGTTTCAGCGAACAGTTTG 1918
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1919 CACAGAGATTCACTTGTTCACCTTAAAGAGCCTTAGAAGCAAAAGAACTAGCTCGGCAG 1978
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluLe 600
Db 1979 CACCTACAGACCACTGGATGAAGTGGAGAAAGAACTAGATCAAAACTACAGGAGATT 2038
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
Db 2039 GATATTTTCAATTAATCAGCTGAAGAACTAAGAGAAATACACAAATAGCAACTCCAG 2098
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIle 640
Db 2099 AAGCAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAGAAATCATTA 2158
QY 641 GluLeuGluLysGlnLysGluAlaGlnArgAlaGlnGluArgAspLysGlnTrp 660
Db 2159 GAATTAAGAAAACAAAAGAAAGAGCCCAAGACGAGCTCAGAAAGGGACAGAGTGG 2218
QY 661 LeuGluHisValGlnGlnAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 2219 CTGGAGCATGTGCAGCGAGGAGCAGCATCAGAGACCAAGAAATCTCCAGAGAGGAA 2278
QY 681 LysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLysGlyLysGlnGlu 700
Db 2279 AAATGAAAAGGAGGAGAGTGTCAAAGAAAGAGTGGCGAGGAAAAGGCAACAGGAA 2338
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal 720
Db 2339 GCACAAAGACAGTGGTGGTGGCTTTTCCATCAACCAACCAAGAACAGCTAAGCCAGCTGTC 2398
QY 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
Db 2399 CAGGCACCTTGGTCCACTGCAGAAAAGGTTCCACTTACCATTTCTGCACAGGAAATGTA 2458
QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
Db 2459 AAAGTGTGTATTAACCGGGCACTGTACCCCTTTGAATCCAGAACCCATGATGAATCACT 2518
QY 761 IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrLeu 780
Db 2519 ATCCAGCCAGGACATAGTCATGGTGGATGAAGCCAACTGGAGAACCCCGCTGGCTT 2578
QY 781 GlyGlyGluLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 800
Db 2579 GGAGGAGAAATTAAGGAAAGACAGAGGTGGTTCCTCTCAAACTATGCAGAGAAATCCCA 2638
QY 801 GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro 820
Db 2639 GAAAAATGAGGTTCCTCGCTCCAGTGAAACCACTGATTCATCAACATCTGCCCTTGGCCCC 2698
QY 821 LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr 840
Db 2699 AAATGGCTTGGGTGAGACCCCGCCCTTTGGCAGTAACCTTTTCCAGAGCCCTCCACG 2758
QY 841 ThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysPro 860
Db 2759 ACCCTTAATTAATGGGCGGACTTCAGCTCCACGTGGCCACCAGCAGCAATGAGAAACCA 2818
QY 861 GluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThrValProSerAlaGly 880
Db 2819 GAAACGGATAAATCGGATGATGGGCGCCCTCTCTCACCGTTCCAAAGTGGCGGC 2878
QY 881 GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro 900

Db 2879 CAGTTAAGCAGAGAGTCCGCTTTACTCCAGCCACGGCCACTGGCTCTCTCCCGTCTCCT 2938
QY 901 ValLeuGlyGlnGlyGluLysValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrpArg 920
Db 2939 GTCTAGGCCAGGGTGAAGAGTGGAGGGGTACAAGCTCAAGCCCTATATCTTGGAGA 2998
QY 921 AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGln 940
Db 2999 GCCAAAAAGACACCACTTAATTTTAAACAAATATGATCATCACCGCTCTGGAAACAG 3058
QY 941 GlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrpPheProLysSerTyr 960
Db 3059 CAAGACATGTGGTGGTTGGAGAGTTCAAGGTCAAGAGGGTTGGTTCCCAAGTCTTAC 3118
QY 961 ValLysLeuIleSerGlyProIleArgLysSerThrSerMetAspSerGlySerSerGlu 980
Db 3119 GTCAAACTCATTTTCAGGGCCCATAGAAGTCTTACAAGCATGATCTTGGTCTTCAGAG 3178
QY 981 SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu 1000
Db 3179 AGTCCTGTAGTCTAAAGCGAGTAGCTCTCCAGCAGCCAGCGGTCTGTTTCGGGAGAA 3238
QY 1000 ----- 1000
Db 3239 GAATTTATGCCATGTACACTTACGAGAGTTCTTGACGAGAGATTTAACTTTCAGCAA 3298
QY 1000 ----- 1000
Db 3299 GGGGATGTGATTTTGGTTTCCAAAGAAAGATGTGACTGTGGTGGACAGGAACAGTGGCGAC 3358
QY 1000 ----- 1000
Db 3359 AAGCCCGAGTCTTCCCTTCTAACTATGTGAGGCTTAAAGATTCAGAGGGCTCTGGAAC 3418
QY 1001 ----- GluIleAlaGlnValIleAlaSerTyr 1009
Db 3419 GCTGGGAAAACAGGGAGTTTAGGAAAACCTGAAATTCGCCAGGTATTGCTCTATAC 3478
QY 1010 ThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuIleLeuIleArgLys 1029
Db 3479 ACCGCCACCGGCCCGAGCAGCTCACTCTCGCCCTGGTCAGCTGATTTTGTATCCGAAA 3538
QY 1030 LysAsnProGlyGlyTyrTrpGluGlyGluLeuGlnAlaArgGlyLysLysArgGlnIle 1049
Db 3539 AAGAACCCAGGTGGTGGTGGAGGAGAGCTGCAAGCACGTGGGAAAAAAGCGCCAGATA 3598
QY 1050 GlyTrpPheProAlaAsnTyrValLysLeuLeuSerProGlyThrSerLysIleThrPro 1069
Db 3599 GGCTGGTTCCAGCTAATTAATGTAAAGCTTAAAGCCCTGGGACGAGCAAAATCCTCCA 3658
QY 1070 ThrGluProProLysSerThrAlaLeuAlaValCysGlnValIleGlyMetTyrAsp 1089
Db 3659 ACAGAGCCACTAAGTCAACAGCATTAGCGCAGTGTGCCAGGTGATTGGGATGTACGAC 3718
QY 1090 TyrThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGlnIleIleAsnValLeu 1109
Db 3719 TACACCGCGAGAAATCAGCATGAGTGGCCCTTCAACAGGGCCAGATCATCAACGTCTC 3778
QY 1110 AsnLysGluAspProAspTrpTyrLysGlyValValAsnGlyGlnValGlyLeuPhePro 1129
Db 3779 AACAGAGAGACCTTGAAGTGGTGGAAAGAGAGTCAATGACACAGTGGGGCTCTTCCCA 3838
QY 1130 SerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
Db 3839 TCCAAATATGTGAAGCTGACACAGACATGGACCAAGCCAG 3880

RESULT 3
AF114488
LOCUS
DEFINITION Homo sapiens intersextin short isoform (ITSN) mRNA, complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824
KEYWORDS

Db 1409 GAGCGCCTGGCCAGCTGGAGCGGGCGGAGCAGAGAGAGAGCGTGGAGCCCGAGGAG 1468
Qy 401 GlnGluArgLysArgGlnLeuGlnLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
Db 1469 CAAGAGCGCAAAAGACAACTGGAACCTGGAGAGCAACTGGAAAGCAGCGGAGCTAGAA 1528
Qy 421 ArgGlnArgGluGlnGluArgArgLysGluLysGluArgArgGluAlaLysArgGlu 440
Db 1529 CGGCGAG 1588
Qy 441 LeuGluArgGlnArgGlnLeuGluArgGlnLeuGluArgGlnLeuGluLeuLeuGln 460
Db 1589 CTTGAAGGCAACGCAACTTGAGTGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1648
Qy 461 ArgAsnLysGluGlnGluAspLysValValLeuLysAlaLysLysLysLysLysLys 480
Db 1649 AGAAACAAAGAAACAGAGGACATAGTTGACTGAAAGCAAGAAAGAAAGAAAGAAAT 1708
Qy 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGluLysLysLysLysLysLys 500
Db 1709 GAATTAGAAGCTCTAATGATATAAAGACATCACTAGAGGGGAAACTTCAAGATATCAGA 1768
Qy 501 CysArgLeuThrThrGlnArgGlnGluLysGlnSerThrAsnLysSerArgGluLeuArg 520
Db 1769 TGTGGATTGACCCCAAGGCAAGAAATGAGAGCAACAACAAATCTAGAGAGTTGAGA 1828
Qy 521 IleAlaGluLysLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGluArg 540
Db 1829 ATTGCCGGAATCACCATCTACAGCAACAATTAAGCAATCTCAGCAAAATGCTTGAAGA 1888
Qy 541 LeuLysProGluLysGlnLysLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1889 CTTATTCCAGAAAACAGATCTCAATGACCAATTAAGCAATTAAGCAATTAAGCAAT 1948
Qy 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1949 CACAGAGATTCACTTGTACCTTAAAGAGCCTTAGAACAAGAACTAGCTCGGCAG 2008
Qy 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluLe 600
Db 2009 CACTACGAGCAACCACTGGATGAAGTGAGAGAAAGAACTAGATCAAACTACAGGAGAT 2068
Qy 601 AspLysPheAsnAsnGlnLeuLysGluLeuArgGluLysHisAsnLysGlnGlnLeuGln 620
Db 2069 GATATTTTCAATTAATCAGCTGAAGGAACCTAAGAGAAATACACAAATAGCAACACTCCAG 2128
Qy 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGlnGluArgLysLysLysLys 640
Db 2129 AAGCAAAAGTCCATGGAGGCTGACGCTGAACAGAAAGCAAGAAAGCAAGAAAGATCAT 2188
Qy 641 GluLeuGluLysGlnLysGluGluAlaGlnArgAlaGlnArgAlaGlnArgAspLysGlnTrp 660
Db 2189 GAATTAGAAAACAAAGAAAGAGAGCCCAAGACGAGCTCAGGAAAGGAGCAAGCAGTGG 2248
Qy 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 2249 CTGGAGCATGTGACGAGGAGGAGCAGAGCATCAGAGACCAAGAAACTCCACGAGAGGAA 2308
Qy 681 LysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLysGlyLysGlnGlu 700
Db 2309 AAATGAAAAGGAGGAGAGTGTCAAAAAGAGATGGCGAGGAAAAGGCAAAAGCAAGGAA 2368
Qy 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGlnProAlaLysProAlaVal 720
Db 2369 GCACAGACAGCTGGGTGGCTTTCATCAACCAACCAAGCAAGCAAGCAAGCAAGCAAG 2428
Qy 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrLysSerAlaGlnGluAsnVal 740
Db 2429 CAGCACCTCGTCCACTGAGAGAAAAGGTCCACTTACCATTTCTGACACGAGAAAATGTA 2488
Qy 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluLysThr 760
Db 2489 AAAGTGTGTATTACCGGGCAGCTGACCCCTTTGATCCAGAGCCATGATGAATCACT 2548

Qy 761 IleGlnProGlyAspLysValMet-----ValAspGluSerGlnThrGly 775
Db 2549 ATCCAGCCAGGAGACATAGTCTGTTAAAGGGGAATGGGTGGATGAAAGCAAACTGGA 2608
Qy 776 GluProGlyTrpLeuGlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyr 795
Db 2609 GAACCCGGCTGGCTTGGAGAGAAATTAAGGAAAGACAGGGTGGTTCCTCGAAACTAT 2668
Qy 796 AlaGluLysLysLeuProGluAsnGluValProAlaProValLysProValThrAspSerThr 815
Db 2669 GCAGAGAAATCCAGAAATGAGGTTCGCGCTCCAGTGAACCAAGTACTGATTCACA 2728
Qy 816 SerAlaProAlaProLysLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSer 835
Db 2729 TCTGCCCTTGGCCCAAACTGGCTTGGCTGAGACCCGCCCTTTGGCAGTAACCTCT 2788
Qy 836 SerGluProSerThrThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSer 855
Db 2789 TCAGAGCCCTCCAGACCCCTATAAATCTGGCGCAGCTTACGCTCCACGTGGCCACAGC 2848
Qy 856 ThrAsnGluLysProGluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThr 875
Db 2849 ACGAATGAGAAACCAAGACGGATACTGGATGCATGGCAGCCGCCCTCTCTCACC 2908
Qy 876 ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly 895
Db 2909 GTTCCAAGTGGCGGCAGTTAAGCAGAGGTCGCGCTTTACTCCAGCACCGCACTGGC 2968
Qy 896 SerSerProSerProValLeuGlyGlnGlyGluLysValGluGlyLeuGlnAlaGlnAla 915
Db 2969 TCTCTCCCGTCTCTGTCTAGGCCAGGGTGAAGGTGGAGGGCTTACAGCTCAAGCC 3028
Qy 916 LeuTyrProTrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValLys 935
Db 3029 CTATATCTTGGAGAGCCAAAAGACCACTTAAATTTTAAACAAATGATGTCATC 3088
Qy 936 ThrValLeuGluGlnGlnAspMetTrpTrpPheGlyGluValGlnGlnLysGlyTrp 955
Db 3089 ACCGTCTCGAAACAGCAAGACATGTGTGTGTTGGAGAGTTCAAGGTTCAGAAAGGTGG 3148
Qy 956 PheProLysSerTrpValLysLeuLysLeuSerGlyProLysArgLysSerThrSerMetAsp 975
Db 3149 TTCCCAAGTCTTACGTGAACCTCACTTTCAGGGCCCAATAAGGAAGTCTCAAGCATGGAT 3208
Qy 976 SerGlySerSerGluSerProAlaSerLeuLysArgValAlaSerProAlaLysPro 995
Db 3209 TCTGTCTTTCAGAGAGTCTGTCTAGTCTAAAGCGAGTAGCTCTCCAGCAGCAAGCCG 3268
Qy 996 ValValSerGlyGlu----- 1000
Db 3269 GTCGTTTTCGGGAGAGAAATTTATGTCATGTACACTTACGAGAGTTCTGAGCAAGGAGAT 3328
Qy 1000 ----- 1000
Db 3329 TTAACCTTTCAGCAAGGGAGTGTATTTGGTTACCAAGAAAGATGTGTACTGGTGACA 3388
Qy 1000 ----- 1000
Db 3389 GGAACAGTGGCCGACAGGCCGGGAGTCTTCCCTTCTAACTATGTGAGGCTTAAAGATTCA 3448
Qy 1001 ----- 1001
Db 3449 GAGGGCTCTGAACTGCTGGGAAAAACAGGGAGTTTAGGAAAAAAACCTGAAATTTGCCAG 3508
Qy 1005 ValLeuLysSerTrpThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeu 1024
Db 3509 GTTATTGCTCTATACACCGCCACCGGCCCGGAGAGCTACTCTCGCCCTTGGTCAAGCTG 3568
Qy 1025 IleLeuLysArgLysLysAsnProGlyGlyTrpTrpGluGlyGluLeuGlnAlaArgGly 1044
Db 3569 ATTTTGTATCGAAAAAAGAAACCCAGGTTGGATGTGGGAGAGAGAGCTGCAAGCAGCTGGG 3628

```
QY 1045 LysLysArgGlnIleGlyTyrPheProAlaAsnTyrValLysLeuLeuSerProGlyThr 1064
Db |||
Db 3629 AAAAGGCCAGATAGCTGGTCCAGCTAAATTATGTAAAGCTTCTAAGCCCTGGAGC 3688
QY 1065 SerLysIleThrProThrGluProProLysSerThrAlaLeuAlaValCysGlnVal 1084
Db |||
Db 3689 AGCAAAATCACTCCAAACAGAGCCACCTAAGTCAACAGCAATAGCGGAGTGTGCCAGTG 3748
QY 1085 IleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGln 1104
Db |||
Db 3749 ATTGGGATGACGACTACCGCCAGAAATGACGATGAGCTGGCTTCAACAAGGGCCAG 3808
QY 1105 IleIleAsnValLeuAsnLysGluAspProAspTyrPheLysGlyGluValAsnGlyGln 1124
Db |||
Db 3809 ATCATCAACGTCCTCAACAGGAGGACCTGACTGGTGGAAAGGAGAGTCAATGGACAA 3868
QY 1125 ValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
Db |||
Db 3869 GTGGGGCTCTCCATCCATTATGTGAAGCTGACCAAGATGGACCATGGACCAAGCCAG 3925

RESULT 4
BD205034 5458 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205034
VERSION BD205034.1 GI:33014804
KEYWORDS JP 2002511267-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5458)
Korenberg,J.R. and Chen,X.N.
AUTHORS Isolated SH3 gene relating to myeloproliferative disorders and
TITLE leukemia and utilization thereof
JOURNAL Patient: JP 2002511267-A 2 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/2
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PI 16-APR-1998 US 60/082007
PR JULIE R KORENBERG,XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

CC leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..5458 /organism='Homo sapiens (human)'
FEATURES
source Location/Qualifiers
1..5458
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 5,34e-229 Length: 5458
Score: 5837.00 Matches: 1143
Percent Similarity: 93.77% Conservative: 0
Best Local Similarity: 93.77% Mismatches: 0
Query Match: 99.02% Indels: 76
DB: 6 Gaps: 2

US-09-720-934-2 (1-1143) x BD205034 (1-5458)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTyrAlaIleThrValGlu 20
|||
```

```
Db 267 ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCCTGGATATCTCGGGCCATAACTGTAGAG 326
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db |||
Db 327 GAAAGACGAGCAGATGATCAGAGTTCCATAGTTTAAAGCCAATATCTGGATTCATTACT 386
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db |||
Db 387 GGTGATCAAGCTAGAAACTTTTTTCAATCTGGGTACCTCAACCTGTTTGTAGACAG 446
QY 61 IleTyrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db |||
Db 447 ATATGGGCACCTAGCTGATCAATGAATGATGGAAGATGGATCAAGTGGAGTTTTCATA 506
QY 81 AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db |||
Db 507 GCTATGAAACTTATCAAACTGAAGCTACAAGGATATCAGTACCCTCTGCACCTCCCT 566
QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db |||
Db 567 GTCATGAAACAGCAACCACTTGTCTATTCTTAGCGCACGACATTTGGTATGGAGGTATC 626
QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db |||
Db 627 GCCAGCATGCCACCGCTTACAGCTGTGTCTCCAGTGGCAATGGATCCATTTCCAGTTGTT 686
QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db |||
Db 687 GGAATGCTCCAACTTAGTATCTTCTGTTCCCAACAGAGCTGTGCCCCCTGGCTTAAC 746
QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db |||
Db 747 GGGGCTCCCTGTTATACAACTCTGCTGCACTTGTCTCATCTCGCAGCCACATGGCA 806
QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db |||
Db 807 AAGAGTTCTTCTCTTTAGTAGATCTGGTCCAGGGTCAACAACATAAATCAAAATCAAAAG 866
QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTyrAlaValProGlnSer 220
Db |||
Db 867 GCACAGTCATTTGATGTGGCAGTGTCCACCACTGGCAGAGTGGGCTGTCTCTCAGTCA 926
QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db |||
Db 927 TCAAGACTGAAATACAGCAATTTATCAATAGTCAATGCAAACTATGAGTGGACACTTA 986
QY 241 ThrGlyProGlnAlaArgThrIleLeuMetClnSerSerLeuProGlnAlaGlnLeuAla 260
Db |||
Db 987 ACAGGTCCCCAAGCAAGAACTATTCTTATGAGTCAAGTTTACCACAGGCTCAGCTGGCT 1046
QY 261 SerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
Db |||
Db 1047 TCATATGGAATCTTTCTGACATTTGATCAAGATGGAACCTTACAGCAGAGAAATTTATC 1106
QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
Db |||
Db 1107 CTGGCAATGCACTCAATTTGATGTAGTATGTCTGGCCAAACCACTGCCACCTGTCTGCTCT 1166
QY 301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
Db |||
Db 1167 CCAGAAATACATTTCCACCTTCTTTTAGAAGAGTTGATCTGGCAGTGTATATCTGTATA 1226
QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuLeuAspGluGlnGln 340
Db |||
Db 1227 AGCTCAACATCTGTAGATCAGAGGCTACCAGAGGAACCCAGTTTTAGAAGATGAACAACAA 1286
QY 341 GlnLeuGlnLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
Db |||
Db 1287 CAATTAGAAAAGAAATTAACCTGTAACTGTTGAAGATGAAGACGGGAGAACTTTGAACGT 1346
QY 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
Db |||
Db 1347 GGCAACTGGAACTGGAGAAACGAAAGCAAGCTCTCTCTGGAAACAGCAGCGCAAGGAGCAG 1406
```

QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluGlnGlu 400
DB 1407 GAGCGCTGGCCAGCTGGAGCGGGCGGAGCAGGAGAGAGAGAGAGAGAGAG 1466
QY 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnGluLysGlnArgGluLeuGlu 420
DB 1467 CAAGACGCGCAAAAGACAATCGGAACCTGGAGAGCAACTCGAAAGCAGCGGAGCTAGAA 1526
QY 421 ArgGlnArgGluGluGluArgLysGluLeuGluArgGluAlaAlaLysArgGlu 440
DB 1527 CGCAG 1586
QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGln 460
DB 1587 CTTGAAAGGCACGACAACTTGAGTGGGAACGGAATCGAAGGCAAGAACTTAAATCAA 1646
QY 461 ArgAsnLysGluGlnGluAspLysLeuValLeuLysAlaLysLysLysLysLeuGluPhe 480
DB 1647 AGAAACAAAGAAACAGAGGACATAGTTGTACTGAAAGCAAGAAAGAACTTTGGAATTT 1706
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGluLysLysLeuGlnAspLysArg 500
DB 1707 GAATTTAGAGCTCTTAATGATAAAAGCATCACTAGAGGGAACTTCAAGATATCAGA 1766
QY 501 CysArgLeuThrThrGlnArgGlnGluLeuGluSerThrAsnLysSerArgGluLeuArg 520
DB 1767 TGTGCGATTGACCAACCAAGGCAAGAAATTTGAGAGCACAACAAATCTAGAGAGTTGAGA 1826
QY 521 IleAlaGluLeuThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGluArg 540
DB 1827 ATTGCCGAAATCAACCATCTACAGCAACAATTAACAGGAATCTCAGCAAAATGCTTTGGAAGA 1886
QY 541 LeuLeuProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
DB 1887 CTTATTCCGAAAGAAACAGATACTCAATGACCAATTAACAGAGTTGAGCAGAAACAGTTTG 1946
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
DB 1947 CACAGAGATTCACTTGTACACTTAAAGAGCGCTTAGAGCAAAAGAACTAGCTCGGCAG 2006
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluLe 600
DB 2007 CACCTAGAGACCAACTGGATGAGTGGAGAAAGAACTAGATCAAAACTACAGGAGATT 2066
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluLeuHisAsnLysGlnGlnLeuGln 620
DB 2067 GATATTTTCAATATCACTGAGGAACTAAGAGAAATACAAATAGCAACAACTCCAG 2126
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleLe 640
DB 2127 AACCAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAAAGAAAGATCATA 2186
QY 641 GluLeuGluLysGlnLysGluGluAlaGlnArgAlaGlnGluArgAspLysGlnTrp 660
DB 2187 GAATTAGAAAGCAAAAGAGAGAGCGCCAAAGACGAGCTCAGGAAGGGGCAAGAGCTGG 2246
QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
DB 2247 CTGAGCATGTGCGCAGGAGAGCGAGCATCAGAGACCAAGAAACTCCAGAGAGAGAA 2306
QY 681 LysLeuLysArgGluGluSerValLysLysLysAspGluGluLysGlyLysGlnGlu 700
DB 2307 AAACCTGAAAGGAGGAGAGGTGTCAAAAGAGAGGATGGCGAGGAGAAAGGCAACAGAA 2366
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal 720
DB 2367 GCACAAGACAGCTGGGTGGCTTTTCCATCAACCAAGAACCAAGCAAGTAAAGCCAGCTGTC 2426
QY 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
DB 2427 CAGGCACCCCTGGTCCACTGCGAAGAAAGAGGTCCACTTACCATTTCTGCACAGGAAATGTA 2486

QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
DB 2487 AAAGTGGTGTATTACCGGCACCTGTACCCCTTTGAATCCAGAGGCATGATGAATCACT 2546
QY 761 IleGlnProGlyAspIleValMet-----ValAspGluSerGlnThrGly 775
DB 2547 ATCCAGCCAGGAGACATAGTCATGTTAAAGGGGAATGGGTGGATGAAGCCAACTGGA 2606
QY 776 GluProGlyTrpLeuGlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyr 795
DB 2607 GAACCCGGCTGGCTTGGAGGAGAAATTAAGAGAAACACAGGGGTGGTTCCCTGCAACTAT 2666
QY 796 AlaGluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThr 815
DB 2667 GCAGAGAAATCCAGAAATGAGTTCCCGCTCCAGTGAACCACTGATTCACA 2726
QY 816 SerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSer 835
DB 2727 TCTGCCCCCTGCCCAAACTGGCTTGGTGAGACCCCGCCCTTTGGCAGTAACCTCT 2786
QY 836 SerGluProSerThrThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSer 855
DB 2787 TCAGAGCCCTCCACGACCCCTAATAACTGGCGGCACTTCAGCTCCACGTGGCCCAACAG 2846
QY 856 ThrAsnGluLysProGluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThr 875
DB 2847 ACGAATGAGAAACAGAAACCGGATAACTGGGATGCATGGGCAGCCCGCCCTCTCTCAC 2906
QY 876 ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly 895
DB 2907 GTTCCAAAGTCCCGCCAGTTAAGCAGAGTCCGCCCTTTACTCCAGCCAGCCGCTGGC 2966
QY 896 SerSerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAla 915
DB 2967 TCCTCCCGCTCTCTGTGTAGCCAGGTGAAAGGTGGAGGGGTACAAAGCTCAAGCC 3026
QY 916 LeuTyrProTrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIle 935
DB 3027 CTATATCTCTGGAGAGCCAAAGAACCAACCTTAAATTTTAAACAAAATGATGTATC 3086
QY 936 ThrValLeuGluGlnGlnAspMetTrpTrpPheGlyGluValGlnGlnLysGlyTrp 955
DB 3087 ACCGTCTTGAACAGACAGACATGTGGTGGTTTGGAGAGTTCAAGTTCAGAGGGTGG 3146
QY 956 PheProLysSerTyrValLysLeuIleSerGlyProIleArgLysSerThrSerMetAsp 975
DB 3147 TTCCCAAGTCTTACGTGAAACTCATTTACGGGCCCATAGGAAGTCTACAGCATGGAT 3206
QY 976 SerGlySerSerGluSerProAlaSerLeuLysArgValAlaSerProAlaAlaLysPro 995
DB 3207 TCTGGTTCTTCAGAGAGTCTCTGTAGTCTAAAGCGAGTAGCCCTCTCCAGCAGCCAGCCG 3266
QY 996 ValValSerGlyGlu----- 1000
DB 3267 GTCGTTCCGAGAGAGAAATTTATTTGCCATGTACACTTACGAGAGTCTTGAGCAGGAGAT 3326
QY 1000 ----- 1000
DB 3327 TTAACCTTTTCAGCAAGGGGATGTGATTTTGGTTACCAGAAAGATGTTGCTGGTGACA 3386
QY 1000 ----- 1000
DB 3387 GGAACAGTGGCGACAAAGCCGGAGTCTTCCCTTTCTAACTATGTGAGGCTTAAAGATTCA 3446
QY 1001 -----GluIleAlaGln 1004
DB 3447 GAGGGCTCTGAACTGCTGGAAAAACAGGAGTTTAGAAAAAAACCTGAAATGGCCAG 3506
QY 1005 ValIleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeu 1024
DB 3507 GTTATTGCTCATACACCGCCACCGGCCCGCAGAGCTCACTCTCGCCCCCTGGTCAAGTGT 3566
QY 1025 IleLeuIleArgLysLysAsnProGlyGlyTrpTrpGluGlyGluLeuGlnAlaArgGly 1044

```
Db 3567 ATTTGATCCGAAAAGAACCCAGGTGGTGGAGGAGTGCAGACAGTGGG 3626
Qy 1045 LysLysArgGlnIleGlyTyrPheProAlaAsnTyrValLysLeuSerProGlyThr 1064
Db 3627 AAAAAGGCCAGATAGGCTGGTCCAGCTAAATTATGTAAAGCTTCTAAAGCCCTGGGACG 3686
Qy 1065 SerLysIleThrProThrGlnProProLysSerThrAlaLeuAlaValCysGlnVal 1084
Db 3687 AGCAAAATCACTCCAACAGAGCCACCTAAGTCAACAGCATTTAGCGGAGTGTCCAGGTG 3746
Qy 1085 IleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGln 1104
Db 3747 ATTGGGATGACGACTACCCGCGCAGATCAAGTCAAGTCAAGTCAAGCGCCAG 3806
Qy 1105 IleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGluValAsnGlyGln 1124
Db 3807 ATCATCAACGTCCTCAACAGGAGGACCCCTGACTGGTGGAAAGGAGAGTCAATGGACAA 3866
Qy 1125 ValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
Db 3867 GTGGGGCTCTCCCATCCATTAATGTGAAGTGCACAGACATGGACCCAGCCAG 3923
```

RESULT 5

```
AF114487
LOCUS AF114487 6439 bp mRNA linear PRI 16-JUL-2002
DEFINITION Homo sapiens intersectin long isoform (ITSN) mRNA, complete cds.
```

```
ACCESSION AF114487
```

```
VERSION AF114487.1 GI:4808822
```

KEYWORDS

```
SOURCE Homo sapiens (human)
```

ORGANISM

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
1 (bases 1 to 6439)
```

```
Authors: Pucharcos, C., Fuentes, J.J., Casas, C., de la Luna, S., Alcantara, S.,
Arbones, M.L., Soriano, E., Estivill, X., and Pritchard, M.
```

```
Alu-splice cloning of human Intersectin (ITSN), a putative
```

```
multivalent binding protein expressed in proliferating and
```

```
differentiating neurons and overexpressed in Down syndrome
```

```
Bur. J. Hum. Genet. 7 (6), 704-712 (1999)
```

```
99415290
```

```
10482960
```

```
2 (bases 1 to 6439)
```

```
Authors: Pucharcos, C., Fuentes, J.J., Pritchard, M. and Estivill, X.
```

```
Direct Submission
```

```
Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
```

```
Cancer Research Institute, L'Hospitalet de Llo., Avia.
```

```
Castelldefels km. 2.7, Barcelona 08907, Spain
```

```
Location/Qualifiers
```

FEATURES

source

```
1. 6439
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/chromosome="21"
```

```
/map="21q22.1-q22.2"
```

```
1. 6439
```

```
/gene="ITSN"
```

```
269..5434
```

```
/codon_start=1
```

```
/product="intersectin long isoform"
```

```
/protein_id="AA029952.1"
```

```
/db_xref="GI:4808823"
```

```
/translation="MAOPPTPGGSLDIWALTVERAKHQDFHSLKPIGSLFITGDOA
```

```
RNFFQSLGPVLAQIALWALADNMNDGMDQVESIAMKLIKILKQGLPSALPPVM
```

```
KQPVLAASSAPFGNGIASMPPLTAAPVPMGSIPIVVGMSFTLVSVPTAAVPLAN
```

```
GAPVPIQLPAPAPATLPKSSFSRSGGSQLNTLQKAQSFDAVSPVPAWAVP
```

```
QSSRLKRLQFNHDKTMSHGLTGPQARTILMQSLPLQAQIASIWNLSIDIDQGLKLA
```

```
EEFLAMHLIDVAMSGQLPVLPPYIPPSFRVRSGSGISVTSSTVDQRLPEEPV
```

```
LEDQQLQKELPVTFFDKENFERGNLEKRRQALKEQKEQKEQLAQLERAEOE
```

```
RKEPERQEQERQLELEKQERLEPEREERERKEPEREAKERLEPEROLEWE
```

```
RNRQELLNRQNKQEDIVLAKKKTLEFLEBALNDKXHQLEKQLDIRCLRTTQRQ
```

ORIGIN

Alignment Scores:

```
Pred. No.: 646e-229 Length: 6439
Score: 5837.00 Matches: 1143
Percent Similarity: 93.77% Conservative: 0
Best Local Similarity: 93.77% Mismatches: 0
Query Match: 99.02% Indels: 76
DB: 9 Gaps: 2
```

```
US-09-720-934-2 (1-1143) x AF114487 (1-6439)
```

```
Qy 1 MetAlaGlnPheProThrPhePheGlySerLeuAspIleTTrpAlaIleThrValGlu 20
Db 269 ATGCTCAGTTTCCAAACACCTTTTGGTGGCAGCTGGATATCTGGGCCATTAACGTAGAG 328
Qy 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 329 GAAAGAGGAGCATGATCAGCAGCTTCCATAGTTTAAAGCCATATCTGGATTCATTACT 388
Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 389 GGTGATCAAGCTAGAAAACCTTTTTCATCTGGTGTACCTCAACCTGTGTTAGACAG 448
Qy 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 449 ATATGGGCACCTAGCTGACATGAATAATGATGGAAGATGGATCAAGTGGAGTTTCCATA 508
Qy 81 AlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 509 GCTATGAACTTATCAAACTGAAGCTACAAGATATCAGCTACCTCTGCATCTCCCTCCT 568
Qy 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db 569 GTCATGAAACAGCAACCACTGCTGCTTCTAGCCACAGCAGCATTTGGTATGGGAGGTATC 628
Qy 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db 629 GCCAGCATGCCACCGCTTACAGCTGTGTCTCAGTGCACCAATGGATTCATTCAGTTGTT 688
Qy 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db 689 GGAATGCTCCAAACCTAGTATCTTCTGTTCCACAGCAGCTGTGCCCCCTCCCTGGCTAAC 748
Qy 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db 749 GGGGCTCCCCCTGTTATACAAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
Qy 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db 809 AAGAGTTCTTCTCTTTAGTAGATCTGGTCCAGGGTCACAACTAAACCACTAAATTAACAAAG 868
```

Qy	201	AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer	220
Db	869	GCAAGTCATTTGATGTGGCCAGTGTCCCAACAGTCGACAGTGGGCTGTTCTCTCACTCA	928
Qy	221	SerArgLeuLysTyArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu	240
Db	929	TCAAGACTGAAATACAGGCAATTAATCAATAGTCATGACAAAACCTATGAGTGACACTTA	988
Qy	241	ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla	260
Db	989	ACAGGTCCTCCCAAGCAAGAACTATTCTTATCGAGTCAGATTACCAAGGCTCAGCTGGCT	1048
Qy	261	SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle	280
Db	1049	TCAATATGGAAATCTTTCTGCAATTCATCAAGATGCAAACTTACAGCAGAGCAATTTATC	1108
Qy	281	LeuAlaMetHisLeuIleAspValAlaMetSerClyGlnProLeuProValLeuPro	300
Db	1109	CTGGCAATGACCTCATTTGATGTAGTATGTGTGGCCAACCACTGCCACCTGCTGCTGCT	1168
Qy	301	ProGluTyIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle	320
Db	1169	CCGAATAATCTCCACCTCTTTATTAGAAGAGTTCGATCTGGCAGTGGTATATCTGTCA	1228
Qy	321	SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln	340
Db	1229	AGCTCAACATCTGTAGATCAGAGCTACAGAGGAACCACTTTTAGAAGATGAACAACAA	1288
Qy	341	GlnLeuGluLysLysLeuProValThrPheGlnAspLysLysArgGluAsnPheGluArg	360
Db	1289	CAATTAGAAAGAAATTACTGTGAACCTTTGAAGATAAGAGCGGGAGAACTTTGAACGT	1348
Qy	361	GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGlnGlnArgLysGluGln	380
Db	1349	GGCAACCTGGAACTGGAGAAACAGAGCAAGCTCTCTCTGGAAACAGACGCAAGGACAG	1408
Qy	381	GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGlnGlu	400
Db	1409	GAGCGCTGGCCCACTGGAGCGGGCGGAGCAGGAGAGGAAGGAGCGTGAGCGCCAGGAG	1468
Qy	401	GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu	420
Db	1469	CAAGAGCGCAAAAGACAACCTGGAACCTGGAGAAGCAACTGGAAGAACAGACGCGAGCTAG	1528
Qy	421	ArgGlnArgGluGluGluArgArgLysGluIleGluArgGluAlaAlaLysArgGlu	440
Db	1529	CGGCAGAGAGAGGAGAGAGAGAAAGAAATTCAGAGCGCAGAGGCTGCAGAAACGGGAA	1588
Qy	441	LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGluLeuLeuAsnGln	460
Db	1589	CTTGAAAGGCAACGACAACTTTGAGTGGAAACGGAATCGAAGGCAAGAACTACTAAATCAA	1648
Qy	461	ArgAsnLysGluGlnGluAspIleValIleLeuLysAlaLysLysThrLeuGluPhe	480
Db	1649	AGAAACAAAGAAACAAAGAGGACATAGTTGTACTGAAGCAAGAAAAAGACTTTGGAATTT	1708
Qy	481	GluLeuGluAlaLeuAsnAspLysHisGlnLeuGluGlyLysLeuGlnAspIleArg	500
Db	1709	GAATTAGAGCTCTAAATGATAAAAGCATCAACTAGAGGGAAACCTCAAGATATCAGA	1768
Qy	501	CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg	520
Db	1769	TGTGCAATGACCAACCAAGCAAGAAATTCAGAGCAACAAACAAATCTAGAGAGTTGAGA	1828
Qy	521	IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArg	540
Db	1829	ATTCCGGAATTCACCCATCTACAGCACAATTAACAGGAATCTCAGCAAAATGCTTTGGA	1888
Qy	541	LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu	560
Db	1889	CTTATTCAGAAAAACAGATACTCAATGACCAATTAACAAAGTTACAGCAGACAGTTTG	1948
Qy	561	HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln	580

Db	1949	CACAGAGATTTCACCTTGTTTACCTTTAAAGAGAGCCTTTAGAACCAAAAGAACTAGTCTGCAG	2008
Qy	581	HISLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle	600
Db	2009	CACCTACGAGACCAACTGGATGAGTGGAGAGAAGAACTAGATCAAACTACAGGAGATT	2068
Qy	601	AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln	620
Db	2069	GATATTTTCAATATATCAGCTGAAGGAACCTAAGAGAAATACACAATAAGCAACAACCTCCAG	2128
Qy	621	LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGlnGlnGluArgLysIleIle	640
Db	2129	AAGCAAAAGTCCATGGAGGCTGAACGACTGAAAACAGAAAGAACCAAGAAAGATCATTA	2188
Qy	641	GluLeuGluLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTirp	660
Db	2189	GAATTAGAAAAACAAAAGAAGAACCCAAAGACGACTCAGAAAGAGGACACAGCAGTGG	2248
Qy	661	LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu	680
Db	2249	CTGGAGCATGTGCAGCAGGAGGACGACATCAGAGACCAAGAAACTCCACGAAGAGGAA	2308
Qy	681	LysLeuLysArgGluGluSerValLysLysAspGlyGluGlnLysGlyLysGlnGlu	700
Db	2309	AAACTGAAAAGGAGGAGAGTGTCAAAAAGAAAGGATGGCGAGGAAAAAGCAACAGGAA	2368
Qy	701	AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal	720
Db	2369	GCACAGACAGCTGGGTGCGCTTTTCCATCAACCAACCAAGAACCGACTAGCCAGCTGTC	2428
Qy	721	GlnAlaProTirpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal	740
Db	2429	CAGCACCCCTGGTCCACTGCAGAAAAAGGTCCCACTTACCATTCTGCACAGGAAAAATGTA	2488
Qy	741	LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr	760
Db	2489	AAAGTGGTGATTATACCGGGCAGCTGTACCCCTTTGAAATCCAGAACCCATGATGAATCACT	2548
Qy	761	IleGlnProGlyAspIleValMet-----ValAspGluSerGlnThrGly	775
Db	2549	ATCCAGCCAGAGACATAGTCATCGTTAAAGGGGAATGGGTGGATGAAGCCAAACTGGA	2608
Qy	776	GluProGlyTirpLeuGlyGlyGluLeuLysGlyLysThrGlyTirpPheProAlaAsnTyr	795
Db	2609	GAACCCGGCTGGCTTTGGAGGAGAATTAAAAGGAAAGACAGGGTGGTTCCCTGCAAACTAT	2668
Qy	796	AlaGluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThr	815
Db	2669	GCAGAGAAATCCACAGAAATGAGGTTCCCGCTCCAGTGAACCAAGTACGTATTCAACA	2728
Qy	816	SerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSer	835
Db	2729	TCTGCCCTTGCCCCAACTGGCTTCGCTGAGACCCCGCCCTTTGGCAGTAACCTCT	2788
Qy	836	SerGluProSerThrThrProAsnAsnTirpAlaAspPheSerSerThrTirpProThrSer	855
Db	2789	TCAGAGCCCTCCACGACCCCTTAATACTGGCCGACTTCAGCTCCACGTGGCCCAACCCAGC	2848
Qy	856	ThrAsnGluLysProGluThrAspAsnTirpAspAlaTirpAlaGlnProSerLeuThr	875
Db	2849	ACGAATGAGAAACACAGAAACCGGATAACTGGGATGTCATGGGCAGCCCAAGCCCTCTCCACC	2908
Qy	876	ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly	895
Db	2909	GTTCCAGTGC CGGCCAGTTAAGGCAGAGGTTCGCGCTTTTACTCCAGCCACGGCCACTGGC	2968
Qy	896	SerSerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAla	915
Db	2969	TCCTCCCGCTCTCTGTGCTAGGCCAGGGTGAAAAGGTGGAGGGGCTACAAGCTCAAGCC	3028
Qy	916	LeuTyrProTirpArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIle	935

Db	107	ATGGCTCAGTTTCCAAACACCTTTGGTGGCAGCCTGGATATCTGGGCCCAATACTGTAGAG	166	Db	1187	GGCAACCTGGAACTGGAGAAACGAAGCAAGCTCTCTCTGGAACAGCAGCGCAAGGAGCAG	1246
Qy	21	GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr	40	Qy	381	GluArgLeuAlaGlnLeuGluArgAlaGluGlnGlnArgLysGluArgGluArgGlnGln	400
Db	167	GAAGAGCGAAGCATGATCAGCAGTTTCCATAGTTTAAAGCCAAATATCTGGATTCAATCT	226	Db	1247	GAGCGCTTGGCCAGCTGGAGCGCGGAGCAGGAGAGGAAGAGCGCTGAGCGCCAGAG	1306
Qy	41	GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln	60	Qy	401	GlnGluArgLysArgGlnLeuGlnLeuGlnLysGlnLeuGluLysGlnArgGluLeuGlu	420
Db	227	GGTGATCAAGCTAGAAACTTTTTTTTCAATCTGGGTACCTCAACCTGTTTAGCACAG	286	Db	1307	CAAGAGCGCAAAAGACAACCTGGAACCTGGAGAGCAACTCGAAAGCAGCGGAGCTAGAA	1366
Qy	61	IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle	80	Qy	421	ArgGlnArgGluGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu	440
Db	287	ATATGGCACTAGCTGACATGATATGATGATGATGATGATGATGATGATGATGATGAT	346	Db	1367	CGCGAG	1426
Qy	81	AlaMetLysLeuLysLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuPro	100	Qy	441	LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGln	460
Db	347	GCTATGAAACTTATCAAACTGAAGCTACAGAGATACAGAGATACAGTACCTTCCCTT	406	Db	1427	CTTGAAGCAACGACCAACTTGAGTGGGAACGGAATCGAAGGCAAGAACTTCTAAATCAA	1486
Qy	101	ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle	120	Qy	461	ArgAsnLysGlnGlnGlnAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe	480
Db	407	GTGATGAACGACCAACAGTTGCTATTTCTAGCGCACCACTTTCGTATGGAGGTATC	466	Db	1487	AGAAACAAAGAACAGAGGACATAGTTGTACTGAAAGCAAGAAAGAAAGACTTTTGAATTT	1546
Qy	121	AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal	140	Qy	481	GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg	500
Db	467	GCAGCATGCCCGCTTACAGCTGTGTCTCCAGTCCCAATGGGATCCATTCAGTTGTT	526	Db	1547	GAAATTAGAAGCTCTAATGATTAAGCATCAACTAGAGGGGAACTTCAAGATATCAGA	1606
Qy	141	GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn	160	Qy	501	CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg	520
Db	527	GGAAATGCTCCAAACCTTAGTATCTTGTTCACAGCAGCTGTGCCCTTGGCTTAAC	586	Db	1607	TGTCGATTGACCACCCCAAGGCAAGAAATTGAGAGCACAACAAATCTAGAGAGTTGAGA	1666
Qy	161	GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaThrLeuPro	180	Qy	521	IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArg	540
Db	587	GGGGCTCCCTGTTATACAACCTGCTGCTGCATTTGCTCATCTGCAGCCACATTTGCCA	646	Db	1667	ATTGCCGAATCACCCATCTACAGCAACAAATTACAGAAATCTCAGCAAAATGCTTGGAGA	1726
Qy	181	LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys	200	Qy	541	LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu	560
Db	647	AGAGATTCTTCTTTAGTAGATCTGGTCCAGGGTCACAATAAACACTAAATTTACAAAAG	706	Db	1727	CTTATTTCCAGAAAAACAGATACTCAATGACCAATTAAACAAGTTTCCAGCAGAAACAGTTTG	1786
Qy	201	AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer	220	Qy	561	HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln	580
Db	707	GCACAGTCATTGATGGCCAGTGTCACACAGTGGCAGTGGCTGTTCCTCAGTCA	766	Db	1787	CACAGAGATTCACTTGTACACTTAAAGAGCCCTTAAAGCAAAAGAACTAGCTCGGCAG	1846
Qy	221	SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu	240	Qy	581	HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle	600
Db	767	TCAAGGCTGAATACAGCAATATTTCATAGTCATGACAAAACATGAGTGGACACTTA	826	Db	1847	CACCTACGAGACCACTGGATGAATGAGTGGAGAAAGAACTAGATCAAACTACAGAGATT	1906
Qy	241	ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla	260	Qy	601	AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln	620
Db	827	ACAGGTCCTCCCAAGCAAGAACTATTCTTATGCAAGTTCACAGGCTCAGCTGGCT	886	Db	1907	GATATTTTCAATTAATCAGCTGAAGAACTAAGAGAAATAACAATAAGCAACACTCCAG	1966
Qy	261	SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle	280	Qy	621	LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGlnGlnArgLysIleIle	640
Db	887	TCAATATGGAAATCTTCTGCAATGATCAAGATGGAAAACTTACAGAGAGAAATTATC	946	Db	1967	AAGCAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAAGAACCAAGATCATA	2026
Qy	281	LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro	300	Qy	641	GluLeuGluLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrp	660
Db	947	CTGGCATGCACCTCATTTGATGATAGTATGTCCTGGCCAAACCACTGCCACCTGTCTGCCT	1006	Db	2027	GAATTTAGAAAAACAAAAAGAAAGCCCAAGACGAGCTCAGAAAGGGCAAGCAGTGG	2086
Qy	301	ProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle	320	Qy	661	LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgProArgLysLeuHisGluGlu	680
Db	1007	CCAGATACATCTCCACCTTCTTTTAGAAGATTCGATCTGGCAGTGTATATCTGTCTATA	1066	Db	2087	CTGGAGCATGTGCAGCAGGAGGAGCAGCATCAGAGACCAAGAAAACTCCACGAAGAGAA	2146
Qy	321	SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGlnGlnGln	340	Qy	681	LysLeuLysArgGluGluSerValLysLysAspGlyGlnGluLysGlyLysGlnGlu	700
Db	1067	AGCTCAACATCTGAGATCAGAGGCTACAGAGAACGAGTTTATAGAGATGAACACAA	1126	Db	2147	AAACTGAAAAGGGAGGAGAGTGTCAAAAGAAAGAGATGGCGAGGAAAAAGGCAACAGAA	2206
Qy	341	GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg	360	Qy	701	AlaGlnAspLysLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal	720
Db	1127	CAATTAGAAAAGAAATTACCTGTAACTGTTTGAAGATTAATAAAGCGGAGAACTTTGAACGT	1186	Db	2207	GCACAAAGCAAGCTGGGTGGCTTTTCCATCAACACCAAGAACCCAGCTTAGCCAGCTGTC	2266
Qy	361	GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln	380	Qy	721	GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal	740
				Db	2267	CAGCACCTTGTTCCACTGCAGAAAAGGTTCCACTTACCATTTCTGCACAGGAAAAATGTA	2326

```
QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluLeuThr 760
Db 2327 AAAGTGGTGTATTACCGGGCACTGTACCCCTTTGAATCCAGAGCCATGATGAATCACT 2386
QY 761 IleGlnProGlyAspIleValMet-----ValAspGluSerGlnThrGly 775
Db 2387 ATCCAGCCAGGAGACATAGTTCATGGTTAAAGGGGAATGGGTGGATGAAAGCCAAACTGA 2446
QY 776 GluProGlyTyrLeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyr 795
Db 2447 GAACCCCGCTGGCTTGGAGGAGAAATTAAAGAGAAAGACAGGGTGGTTCCCTGCCAACTAT 2506
QY 796 AlaGluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThr 815
Db 2507 GCAGAGAAATCCAGAAATAGAGTTCCTCCGCTCAAGTGAACCAAGTCACTGATTCACA 2566
QY 816 SerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSer 835
Db 2567 TCTGCCCTTGCCTCCCAACTGGCTTGGCTGAGACCCCGCCCTTTGGCAGTAACTCT 2626
QY 836 SerGluProSerThrThrProAsnAsnTyrPheAlaAspPheSerSerThrTyrProThrSer 855
Db 2627 TCAGAGCCCTCCAGACCCCTTAATACTGGGCCGACTTCAGCTCCAGTGGCCCAACGAGC 2686
QY 856 ThrAsnGluLysProGluThrAspAsnTyrAspAlaTyrAlaAlaGlnProSerLeuThr 875
Db 2687 ACGAATGAGAAACCAAGAACGATTAACCTGGGATGATGGGAGCCAGCCCTCTCTCACC 2746
QY 876 ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly 895
Db 2747 GTTCCAAGTGGCGCCAGTTAAGGCAGAGGTCCGCTTTACTCCAGCCAGCGCCACTGGC 2806
QY 896 SerSerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAla 915
Db 2807 TCCTCCCGCTCTCTGTCTAGGCGAGGTGAAAGGTGGAGGGCTACAGCTCAAGCC 2866
QY 916 LeuTyrProTyrArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIle 935
Db 2867 CTATATCTTGGAGAGCCAAAGAGACCAACCTTAATTTACAAAAATGATGTCACTC 2926
QY 936 ThrValLeuGluGlnAspMetTyrTyrPheGlyGluValGlnGlyLysGlyTyr 955
Db 2927 ACCGTCTCGAACAGCAAGCATGTGGTGGTTTGGAGAGTTCAAGGTCAAGGGTTGG 2986
QY 956 PheProLysSerTyrValLysLeuLysSerGlyProIleArgLysSerThrSerMetAsp 975
Db 2987 TTCCCAAGTCTTACGTGAACCTCATTTTCAGGGCCCATAGGAAGTCTACAGCATGGAT 3046
QY 976 SerGlySerSerGluSerProAlaSerLeuLysArgValAlaSerProAlaAlaLysPro 995
Db 3047 TCTGGTTCTTCAGAGAGTCTCTAGTCTAAAGCGAGTAGCTCTCCAGCAGCAAGCCG 3106
QY 996 ValValSerGlyGlu----- 1000
Db 3107 GTCTGTTTCGGGAGAGAAATTTATTGCCATGTATACGAGAGTTCTGAGCAAGAGAT 3166
QY 1000 ----- 1000
Db 3167 TTAACCTTTTCAGCAAGGGATGTGATTTTGGTTACCAAGAAAGATGGTACTGGTGACA 3226
QY 1000 ----- 1000
Db 3227 GGAACAGTGGCGACAAGCGCGAGTCTCTCCCTTCTAACTATGTGAGGCTTAAAGATTCA 3286
QY 1001 -----GluIleAlaGln 1004
Db 3287 GAGGGCTCTGAACTGCTGGGAAACAGGAGATTAGCAAAAAAACCCTGAATTTGCCAG 3346
QY 1005 ValIleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeu 1024
Db 3347 GTTATTGCCTCATACACCGCACCGCGCCGAGGAGTCACTCTCGCCCTTGGTCAGCTG 3406
```

```
1025 IleLeuIleArgLysLysAsnProGlyGlyTyrTyrGluGlyGluLeuGlnAlaArgGly 1044
Db 3407 ATTTTGATCCGAAAAGAACCCAGGTGGATGGTGGGAGGAGAGCTGCAAGCAGCTGGG 3466
QY 1045 LysLysArgGlnIleGlyTyrPheProAlaAsnTyrValLysLeuLeuSerProGlyThr 1064
Db 3467 AAAAAGCGCCAGATAGGCTGGTCCAGCTAATTTATGTATAAGCTTCTAAACCTGGGAGC 3526
QY 1065 SerLysIleThrProThrGluProProLysSerThrAlaLeuAlaAlaValCysGlnVal 1084
Db 3527 AGCAAAATCACTCCAAAGAGAGCCACCTAAGTCAACAGCATTTAGCGGAGTGTGCCAGGTG 3586
QY 1085 IleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLysGlyGln 1104
Db 3587 ATTGGGATGTACGACTACACCGCGCAGATGACCATGAGCTTGGCTTCAACAGGGCCAG 3646
QY 1105 IleIleAsnValLeuAsnLysGluAspProAspTyrTyrLysGlyGluValAsnGlyGln 1124
Db 3647 ATCATCAACGTCTTCAACAGGAGGAGCCCTGACTGGTGGAAAGGAGAGTCAATGGACAA 3706
QY 1125 ValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
Db 3707 GTGGGGCTCTTCCCATCAATTATGTGAAGCTGACCAAGACATGGACCCCAAGCCAG 3763
```

RESULT 7
AF064244 7247 bp mRNA linear PRI 21-NOV-1998
LOCUS Homo sapiens intersectin long form mRNA, complete cds.
DEFINITION AF064244
ACCESSION AF064244
VERSION AF064244.1 GI:3859854
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7247)
AUTHORS Guipponi M., Scott H.S., Chen H., Schebesta A., Rossier C. and Antonarakis S.E.
TITLE Two isoforms of a human intersectin (ITSN) protein are produced by brain-specific alternative splicing in a stop codon
JOURNAL Genomics 53 (3), 369-376 (1998)
MEDLINE 99017974
PubMed 9799604
REFERENCE 2 (bases 1 to 7247)
AUTHORS Guipponi M., Scott H.S., Chen H., Schebesta A., Rossier C. and Antonarakis S.E.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue Michel-Servet, Geneva 4 CH-1211, Switzerland
FEATURES
source Location/Qualifiers
1..7247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
107..5272
/codon_start=1
/product="intersectin long form"
/protein_id="AAC78611.1"
/db_xref="GI:3859854"
/translations="MAQFPPTFGGSLDIWALTVEERAKHDQFHSLKPIGSGITGDOA
RNFPSGGLPQPVLAQIWLADNMNDGRMDQVFSIAMKLIKLLQYQLPSALPPVYM
KQPVAISSAPFGMGIAINPPLTAVAPVPMGSIPIVPMSPILVSVPTAAVPLAN
GAPVLPQLPFAHPAATLPKSSFSRSGSGSQNTKIQAKQSFDAVSFVPAEWAVP
QSRRLKYRQFNHSDKTMGHLTPQPARTILMQSLPQAKLASIWNLSIDIDQDKLTA
EFTILAMHLIDVAMSGQPLPVLPPYIPPSFRVRSVSGSISVTSRQRLPEEPV
LEDEQQLEKLPVTPEDKKNFENRGNLEKRRQALLESQKQKQRLAQLERAEOE
RKEREQERKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQ
RNRQELLNQRKEQEDIVLKAKKTKLEFELEALNDKKHQLSGKLODKLDTQRLTQ
EISTNKSRLRIAEITHLQQQLQESQQMLGRLIPEKQILNDQLKQVQQNSLRHDSLV

TLKRALAEKELARQHRLDQDDEVEKETRSKQLQEDIFNNQLKELREIHNKQLOKQKS
 MEAERLKQEQERKITELEKQKEEAQRAERDKQWLEHVQOEDERHQRPLKHEHEEKL
 KREESVKKQDEKSGHQAEDDKLGRFHQHPAKPAVOAPWSTAEKGLKGTWPPA
 KVVYVRLYPPFSRQDEITIOGPDIVMVUKGEWDESOTGEPGLGELGKGTWPPA
 NYAEKIPENEVPAPKVPYDTSAPAPKALARETRAPLAVTSSEPTTNNWADFSST
 WFTSTNEKETPDNDWAQAQSLUTVPSAGQLKQRSAPTEPATATGSSPSPVLGGQEKVE
 GLOAQLXPWRACKNDHNFNDVITVLEQQDMWFWGVEQQQKGFPPKSYVKLISGP
 IRKSTDMDSGSSSEPKARVASPAAPKPVVSGEEFTAMTYTESSEGDLTFQOQGVIL
 VTKKDMDWTGTVGDKAGVFPSPYVRLKDSGSGTAGKTSGLGKPEIAQVIASYTAT
 GEQJTLAQQILILIRKKNPGGWMEGLOARGKKROIOWFPANVVKLLNPGTSKITPT
 EEPKTAALAAQCQVIGMYDYTAGDDELAFNPKQILINVLNKEDPDMWKEVNGQVGLF
 PSNYKLTITMDPQSQWSDLLHMLTPTFRKQGIIEHLIVTEENYVNDLQVTEI
 FQKPLMESLLEKEVAMIFVNMKELIMCNIKLLKALRVKXWGSKMPVKMIGDILS
 AOLPHMQPIRFCSRLNGAALTQOKTDEAPDFKVFVRLKLEMDPRCKMPLSSFILP
 MORVTRYPLI IKNILENTPENHPDHLKHALEKABELCSQVNEGVEKENSRLLEWI
 OAHVOCEGLSEQLVFNSTNCLGPKFLHSGKLYKAKNNKELGFUFPNDFLLLTQTK
 PIGSSGTDKVPSPKSNLOKMYKTPIFLNEVLVKLPDPSGDEPIPHISHIDRVYTLR
 AESINERTAVVQKKAASELYIETEKKREKAYLRSQRATGIGRLMVNVVEGIELKP
 CRSHGKSNPYCEVTMGSQCHI TKTIQDTLNPKNWSNCOFFIROLEQEVLCITVFERDQ
 FSPDDPLGRITRIVADIKKQQSGKGPVTKLLLHEVPTGEBIVVRLDLQLFDEP"
 167. .406
 misc_feature /note="encodes EH domain"
 767. .1936
 misc_feature /note="encodes EH domain"
 2324. .2524
 misc_feature /note="encodes SH3 domain"
 2843. .3019
 misc_feature /note="encodes SH3 domain"
 3110. .3286
 misc_feature /note="encodes SH3 domain"
 3326. .3520
 misc_feature /note="encodes SH3 domain"
 3569. .3748
 misc_feature /note="encodes SH3 domain"
 3836. .4390
 misc_feature /note="encodes GDS domain"
 4649. .4819
 misc_feature /note="encodes PH domain"
 4895. .5143
 misc_feature /note="encodes C2 domain"

ORIGIN

Alignment Scores:

Pred. No.: 1.57e-228 Length: 7247
 Score: 5829.00 Matches: 1141
 Percent Similarity: 93.68% Conservative: 1
 Best Local Similarity: 93.60% Mismatches: 1
 Query Match: 98.88% Indels: 76
 DB: 9 Gaps: 2

US-09-720-934-2 (1-1143) x AF064244 (1-7247)

QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
 Db 107 ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCTCGATATCTGGCCATAACTGTAGAG 166
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 167 GAAGAGCGAAGCATGATCAGCAGTTCATAGTTTAAAGCCAAATATCTGGATTCTACT 226
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 227 GGTGATCAAGCTAGAACTTTTTCATCTGGGTATCCTCAACCTGTTTATAGCACAG 286
 QY 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 287 ATATGGGCACCTAGCTGATGAATTAATGATGAAGAATGGAATCAAGTGGAGTTTTCATA 346
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 347 GCTATGAACCTTATCAACTGAAGCTACAGGATATCAGCTACCTCTGCACTTCCCCCT 406
 QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaPheGlyMetGlyGlyIle 120
 Db 407 GTCATGAAAACAGCAACCATGCTGCTATTTCTAGCGCACCACTTTGGTATGGAGGTATC 466
 QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 467 GCCAGAGATGCCACCGCTTACAGCTGTGCTCCAGTGCCTAATGGATCCATTCAGTTGTT 526
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 Db 527 GGAATGCTCCAAACCCCTAGTATCTTCTGTTCCACAGCAGCTGTGCCCCCTCGGTAAAC 586
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 Db 587 GGGGCTCCCCCTGTTATACAACCTTGCCTGCACTTTGCTCATCTCGCAGCCACATGGCCA 646
 QY 181 LysSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 Db 647 AAGAGTCTCTCTTTTAGTAGATCTGGTCCAGGGTCAACAATAAACAATAAATTAACAAG 706
 QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
 Db 707 GCACAGTCATTTGATGTGCCAGTGTCCACACAGTGGCAGAGTGGGCTGTTCCTCAGTCA 766
 QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 Db 767 TCAAGGCTGAATACAGGCAATTTATTCAATAGTCATGACAAAACTATGAGTGGACACTTA 826
 QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
 Db 827 ACAGGTCCCCAACAGCAAGAACTATTTTATGCACTCAAGTTTACCACAGGCTCAGTGGCT 886
 QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
 Db 887 TCAATATGGAATCTTCTGACATTGATCAAGATGGAACCTTACAGCAGAGGAATTTATC 946
 QY 281 LeuAlaMetHisLeuLeuAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
 Db 947 CTGGCAATGCACCTCATTTGATGATGCTATGCTGGCCAACTGTCGCCACTGCTGCTGCT 1006
 QY 301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
 Db 1007 CAGATAACATTCACCTTCTTTTAGNAGATTCGATCTGGCAGTGGTATATCTGTCATA 1066
 QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
 Db 1067 AGCTCAACATCTCTAGATCAGAGGCTACCAAGAGAACCACTTTTAGAAGATGAACACAA 1126
 QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
 Db 1127 CAATTAGAAAAGAAATTTACCTGTAACGTTTGAAGATAAAAGCGGAGAACTTTTGAACGT 1186
 QY 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnArgLysGluGln 380
 Db 1187 GGCAACCTGGAATCTGGAGAAACGAAGGCAAGCTCTCTGGAAACAGCAGCGCAAGAGCAG 1246
 QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGluGlu 400
 Db 1247 GAGCGCTGGCCACGCTGGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
 QY 401 GlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
 Db 1307 CAGAGCGCAAAAGACAACTGGAACTGGAGAACCTGGAGAACCTGGAAAGAGCGCGGAGTAA 1366
 QY 421 ArgGlnArgGluGluGluArgArgLysGluIleGluArgArgGluAlaIleLysArgGlu 440
 Db 1367 CGGCAG 1426
 QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGln 460
 Db 1427 CTTGAAAGGCAACGCAACTTGAGTGGAAACGGAATCGAAGGCAAGAACTACTAAATCAA 1486
 QY 461 ArgAsnLysGluGlnGluAspIleValLeuLysAlaLysLysThrLeuGluPhe 480

1487 AGAAACAAAGCAAGAGGACATAGTTGTACTGAAAGCAAGAAAGAAAGACTTTTGAATTT 1546
1488
481 GluLeuGluAlaLeuAenAspLysLysHiscInLeuGluGlyLysLeuGlnAspIleArg 500
1547 GAATTAGAGCTCTAAATGATTAAGAGCATCACTAGAGGAAACCTTCAAGATATCAGA 1606
1548
501 CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAenLysSerArgGluLeuArg 520
1549
1607 TGTCGATGACCCACCCAAAGCAAGAAATTTGAGAGCAACAACAATCTAGAGAGTTGAGA 1666
1608
521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnMetLeuGlyArg 540
1609
1667 ATTGCCGAATACCCACTTACAGCAACAATTTACAGGAATCTCAGCAAAATGCTTTGGAAGA 1726
1670
541 LeuIleProGluLysGlnIleLeuAenAspGlnLeuLysGlnValGlnGlnAenSerLeu 560
1671
1727 CTTATTCCAGAAACACAGATCACTCAATGACCAATTTAAACAAAGTTTCAGCAGAAAGTTTG 1786
1728
561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
1729
1787 CACAGAGATTCACCTTGTTACACTTAAAGAGCCTTAGAGCAAAAGAACTAGCTCGGCAG 1846
1788
581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
1789
1847 CACCTACGAGACCAACTGGATGAAGTGAGAAAGAACTAGATCAAAACTACAGAGATT 1906
1848
601 AspilePheAenAenGlnLeuLysGluLeuArgGluIleHisAenLysGlnGlnLeuGln 620
1849
1907 GATATTTTCAATATCAGCTGAAGGAACCTAAGAGAAATACAAATAAGCAACAACTCCAG 1966
1908
621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIle 640
1909
1967 AAGCAAAAGTCCATGGAGCTGACAGCTGAAAACAAGAAACAAAGAAAGATCATA 2026
1970
641 GluLeuGluLysGlnLysGluGluAlaGlnArgAlaGlnGluArgAspLysGlnTrp 660
1971
2027 GAATTAGAAAACAAAAGAAAGAGCCCAAGACAGAGCTCAGGAAAGGGAACAAGCAGTGG 2086
1972
661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
1973
2087 CTGAGAGATGTCAGAGAGGAGGAGAGCATCAGAGACCAAGAAACCTCCAGAAAGAGAA 2146
1974
681 LysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLysGlyLysGlnGlu 700
1975
2147 AAACCTGAAAGGGAGGAGAGTGTCAAAAGAGAGGTGCGGAGGAAAGAAAGGCAACAGAA 2206
1976
701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGlnProAlaLysProAlaVal 720
1977
2207 GCACAAGACAGCTGGTGGTGGCTTTTCCATCAACACCAAGAACCCAGCTTAAGCCAGTGT 2266
1978
721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAenVal 740
1979
2267 CAGGACCCCTGGTCCACTGCAAGAAAGGTCCATTCACATTTCTGCACAGAAAGTGA 2326
1980
741 LysValValTyrtArgAlaLeuTyrtProPheGluSerArgSerHisAspGluIleThr 760
1981
2327 AAAGTGGTGATTAACCGGCACTGTACCCCTTTGAATCCAGAGCAATGATGAATCACT 2386
1982
761 IleGlnProGlyAspIleValMet-----ValAspGluSerGlnThrGly 775
1983
2387 ATCCAGCCAGAGACATAGTCATGTTAAAGGGGAATGGGTGGATGAAAGCAAACTGGA 2446
1984
776 GluProGlyTyrLeuGlyGlyLeuLysGlyLysThrGlyTyrPheProAlaAenTyrt 795
1985
2447 GAACCCGCTGGCTTGGAGAGGAATTAAGAGAAAGACAGAGGTGGTTCCTTGCAACTAT 2506
1986
796 AlaGluLysIleProGluAenGluValProAlaProValLysProValThrAspSerThr 815
1987
2507 GCAGAGAAATCCCAAGAAATCAGGTTCCTCCGCTCAGTGAACCAAGTGAATTCACA 2566
1988
816 SerAlaProAlaProLysLeuAlaLeuArgGlnThrProAlaProLeuAlaValThrSer 835
1989
2567 TCTGCCCTGCCCCCAAACTGGCTTGGTGGAGACCCCGCCCTTTTGGCAGTAACTCT 2626
1990
836 SerGluProSerThrThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSer 855
1991
2627 TCAGAGCCCTCCAGACCCCTTAATCTAGTGGCCGACTTCAGTCCAGCTGCGCCACCCAGC 2686
1992
856 ThrAsnGluLysProGluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThr 875
1993
2687 ACGAATCAGAAACCCAGAAACCGATAAATCTGGATCATGGGAGCCAGCCAGCCCTCTCTCACC 2746
1994
876 ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly 895
1995
2747 GTTCCAAAGTCCCGCCAGTTAAGCAGAGGTCCGCTTTATCTCCAGCCACCGCCAGTGGC 2806
1996
896 SerSerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAla 915
1997
2807 TCTTCCCGCTCTCTGTCTAGCCAGGTGAAAAGGTGGAGGGCTTACAGCTCAAGCTCAAGCC 2866
1998
916 LeuTyrtProTrpArgAlaLysLysAspAsnHisLeuAenPheAenLysAsnAspValIle 935
1999
2867 CTATATCTTTGGAGAGCCAAAAGAACCAACCACTTAAATTTTAAACAAAATGATGTCTATC 2926
1999
936 ThrValLeuGlnGlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrp 955
1999
2927 ACCGTCTCGAAACAGCAAGACATGTGTGTGTGGAGAGTTCAGGTTCAGAGGGTGG 2986
1999
956 PheProLysSerTyrtValLysLeuIleSerGlyProIleArgLysSerThrSerMetAsp 975
1999
2987 TTCCCCAAGTCTTACGTGAACCTCATTTCAGGGCCCATAGGAAGTCTACAGCATGGAT 3046
1999
976 SerGlySerSerGluSerProAlaSerLeuLysArgValAlaSerProAlaAlaLysPro 995
1999
3047 TCTGGTCTTTCAGAGAGTCTCTGTAGTCTAAAGCGAGTAGCTCTCCAGCAGCAAGCCG 3106
1999
996 ValValSerGlyGlu----- 1000
1999
3107 GTCGTTTGGGAGAGAATTTATTGCCATGTACACTTACAGAGATTCTGAGCAAGAGAT 3166
1999
1000 ----- 1000
1999
3167 TTAACCTTTCAGCAAGGGATGTATTTGGTTTACCAGAAAGATGGTACTGGTGACA 3226
1999
1000 ----- 1000
1999
3227 GGAACAGTGGCGCAAGAGCGGAGTCTTCCCTTCTAACTATGTGAGCTTTAAAGATTCA 3286
1999
1001 -----GluIleAlaGln 1004
1999
3287 GAGGCTCTGAACTGTGGGAAACAGGGAGTGTAGGAAAAAACCTGAAATTTGCCGAG 3346
1999
1005 ValIleAlaSerTyrtThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeu 1024
1999
3347 GTTATTGCTCATACACCGCCACCGCCCGGAGCAGCTCACTCTCGCCCTTGGTCAGCTG 3406
1999
1025 IleLeuIleArgLysLysAsnProGlyGlyTyrtTrpGluGlyGluLeuGlnAlaArgGly 1044
1999
3407 ATTTTGTATCCGAAAAAGAACCCAGTGGATGGTGGGAGAGAGAGTGCAGACAGCTGGG 3466
1999
1045 LysLysArgGlnIleGlyTyrtPheProAlaAenTyrtValLysLeuLeuSerProGlyThr 1064
1999
3467 AAAAGCGCCAGATAGGTGTTTCCAGCTTAATTTATGTAAAGCTTCTTAAACCTTGGGAGC 3526
1999
1065 SerLysIleThrProThrGluProProLysSerThrAlaLeuAlaValCysGlnVal 1084
1999
3527 AGCAAAATCACTCCCAACAGAGCCACCTAAGTCAACAGCATTTAGCGGAGTGTGCGAGGTG 3586
1999
1085 IleGlyMetTyrtAspTyrtThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGln 1104
1999
3587 ATTGGGATGTACGACTACCGCGAGAAATGACCATGAGCTTGGCTTCCAAAGGCGCAG 3646
1999
1105 IleIleAsnValLeuAenLysGluAspProAspTrpTrpLysGlyGluValAsnGlyGln 1124
1999
3647 ATCATCAACGCTCTCAACAGGAGGACCTGACTGGTGGAGAAAGAGAGTCAATGAGCAA 3706
1999

QY 1125 ValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
 Db 3707 GTGGGGCTCTCCCATCAATTATGTAGAGCTGACACAGACATGAGCCAGCCAG 3763

RESULT 8
 AF132672
 LOCUS 3812 bp mRNA linear ROD 28-JUN-1999
 DEFINITION Rattus norvegicus EH-domain/SH3-domain containing protein mRNA,
 complete cds.
 ACCESSION AF132672
 VERSION AF132672.1 GI:4838525
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 3812)
 AUTHORS Okamoto,M., Schoch,S. and Sudhof,T.C.
 TITLE EHS1/intersectin, a protein that contains EH and SH3 domains and
 binds to dynamin and SNAP-25. A protein connection between
 exocytosis and endocytosis?
 J. Biol. Chem. 274 (26), 18446-18454 (1999)
 JOURNAL 99303609
 MEDLINE 10373452
 PUBMED

REFERENCE 2 (bases 1 to 3812)
 AUTHORS Okamoto,M., Schoch,S. and Sudhof,T.C.
 TITLE Direct Submission
 JOURNAL Submitted (02-MAR-1999) Center for Basic Neuroscience and HHMI,
 Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
 75235, USA

FEATURES
 source
 1. .3812
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 41..3481
 /note="EHS1"

CDS
 /codon_start=1
 /product="EH-domain/SH3-domain containing protein"
 /protein_id="AAD31026.1"
 /db_xref="GI:4838526"

1
 /translation="MAQPTTFPGSLDIWALTVEERAKHQOQSLKPISGFTIGDQA
 RNFFQSGLPQVLAQIWAALADNKGDMQVEFSIAMKLKLGQYQLPPLPPVM
 KQPAALISSAPAFGIGMAGMPPPLTAVPVPMGSI PVGMSPLVSSVPQAAVPPLAN
 GAPPVQIPLPAFHPAATLPKSSFSRSGFQSLNTKLQAKQAFDVAAPAAAEWAVP
 QSRLLKYROLFNSHDHTMSGLTGPQARTILMOSLSLPAQALASTWNLSDIDQGLTA
 EFTLAMHLIDVAMSGQLPPVLPPEVI PPSFRVRSVSGMSVSISSSADQRLPEEPS
 SEDEQVEKLPVTFEDKKENFERGNLEKRRQALLEGQREKQERLAQLERAQER
 KERQEQERKRLQLEKQLEKQERLEKQERLEKQERLEKQERLEKQERLEKQER
 NRRQELLTORNKQEGIVLVILKARRKTLFELEALNDKKHQLGKLDIRCLATQRE
 IESTNSRELRIAEITHLQOQLOESQMLGLIPEKQILSDQLKQVQNSLHRDSLIT
 LKRALEAKELARQOLDEVEKETSRLQLEIDVFNQLKEIREIHSKQOLQKRSI
 EERLQKQERKSLQLEKQERQVRODKQORHVOEQEORPRKPEEDKLKR
 EDVYKKEAERAKPEVDQKQSLFHEHPQAPKAQAPWTTEKPLTISAQESAKV
 YKRALYFPERSHDEITIQGDIVVMKGEWVDESQTGEPGMLGEPKGTGWFPANTA
 EKIPENISPTAPKPDVLTSAAPKALRETPAPLPTVSEPTTPNNWADFSTFWS
 STNEKPTDNDWMAQPSLTVPAGLQRFSAFTPATATGSSPSVPGQEKVEGLQ
 AQLYPRWAKNDHNLNENKSDVITVLQDQMWVFEVQOGKWPFPKYVLISQFVK
 STSDTCPTPEAPSLKRVSPAAPKPAIPGEILAOVIASVATGPEQLTAPQGLILIR
 KNPFGWBEELQARKKRLQGWFPANVYKLSGTSKIPTFELPKTAVQPAVCQVIG
 "MYDITAQNDDELAFSKGQILNVLSGQVDMWKGEVSGQVLPNSYVYKLTITDMPDQO
 "

ALIGNMENT SCORES:
 Pred. No.: 8,13e-214 Length: 3812
 Score: 5460.00 Matches: 1058
 Percent Similarity: 95.21% Conservative: 35
 Best Local Similarity: 92.16% Mismatches: 47
 Query Match: 92.62% Indels: 8
 DB: 10 Gaps: 4

ORIGIN

US-09-720-934-2 (1-1143) x AF132672 (1-3812)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
 Db 41 ATGGCTCAGTTTCCGACACCTTTTGTGGGAGCTGGACATCTGGGCCATACCGTGGAG 100

QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 101 GAAAGAGCCCAAGCATGACCAAGCAGTTTCCAGAGTTTGAAGCCGATATCTGGATTTCATCACT 160

QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 161 GGTGATCAAGCAAGAACTTCTTTTCCATCTGGGTACTCAGCCCGTCTTAGCACAA 220

QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 221 ATATGGCGCTGGCTGACATGAATAAGGACGGAAGGATGATCAGGTGGAGTTTTCATATA 280

QY 81 AlaMetLysLeuLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 281 GCCATGAAGCTCATCAAACTGAAAGCTGCAAGGATATCAGCTCCCCCGCAGCTTCCCCCT 340

QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyLe 120
 Db 341 GTATGAAGCAGCAGCAGCGGCCCATCTCTAGTGCACCGCTTGTGTATAGGAGGATG 400

QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 401 GCTGGAATGCCACCACTGACAGCTGTTGCTCCCGTGCAATGGCTCCATCCAGTTGTT 460

QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 Db 461 GGAATGTCTCGCCCTTAGTATCTCTGTCCCTCAAGCAGCAGTCCCTCCCTCGCTAAC 520

QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 Db 521 GGGGCCCCCTCTGTGTCATACAGCCTCTGCTGCATTTGCTCATCTCTGCGACACATTGCCA 580

QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 Db 581 AAGAGTTCTTCTTCCAGCAGATCTGTGTCAGGGTCACAATTTAAACATTAACACAGAAG 640

QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
 Db 641 GCACAAATCATTTGATGTAGCCAGCGCCCTCGACGCGCAGAAATGGGTGTGCTCAGTCG 700

QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 Db 701 TCAAGACTGAAATACAGCAGTATTTCACAGTCACGACAGACCATGAGTGGACACTTA 760

QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
 Db 761 ACAGGTCCCCAGGCAAGAACTATTCTCATGCAATCAAGTTTACCAGGCTCAGCTGCT 820

QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
 Db 821 TCAATCTGGAATCTTCCGACATTTGATCAAGATGGAAGCTCACCGCAGAGAAATTCATC 880

QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
 Db 881 CTAGCGATGACCTGATTTGATTTGCTATGCTGTGCTGAGCCACTGCGCGCTGCTGCTGCT 940

QY 301 ProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle 320
 Db 941 CCAGAAATACATCCCTCTCTTCTTAGAAGATTCGCTCTGCGAGTGGGATGTCGTCATATA 1000

QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
 Db 1001 AGCTTTCTGCTGTCAGACACCGCGCTCCCGGAGAGGCATCGTCAGAGGATGAG--CAG 1057

QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
 Db 1058 CAGTGGAAAAGAAGCTGCTGCTGTGACATTTGAAGATTAAGAAGCGGGAGAACTTCGAGCGA 1117

```
Qy 361 GlyAsnLeuGluLeuGluLeuArgGlnAlaLeuLeuGluGlnArgLysGluGln 380
Db 1118 GCGAACCTGGAGCTGGAGAGCGCAGGAGCGCCCTCTCGAGCAGCAGCGCAAGAGCAG 1177
Qy 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluGlnGlu 400
Db 1178 GAGCGCTGGCTCAGCTGGAGCGCGCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
Qy 401 GlnGluArgLysArgGlnLeuGluLeuGluLeuLysGlnLeuGluLysGlnArgGluGlu 420
Db 1238 CAGGAGCGCAAGAGACAGCTAGAGCTGGAGAGCAGCTAGAGAGAGAGAGAGAGAGAGAG 1297
Qy 421 ArgGlnArgGluGluArgLysGlnLeuGluArgGlnGluArgGlnAlaAlaLysArgGlu 440
Db 1298 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
Qy 441 LeuGluArgGlnArgGlnLeuGluArgGlnArgGlnGlnGlnGlnGlnGlnGlnGln 460
Db 1358 CTCGAACGGCAGCAGCTTGAATGGAAACGGAACCGGAGAGAGAGAGAGAGAGAGAG 1417
Qy 461 ArgAsnLysGluGlnGluAspLysValValLysLysLysLysLysLysLysLysLys 480
Db 1418 AGGAACAAGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1477
Qy 481 GluLeuGluAlaLeuAsnAspLysLysLysGlnLeuGluGluGluLysLysLysLys 500
Db 1478 GAATTAAGAGCTCTGAATGACAAAGAGATCAGCTGGAGAGAGAGAGAGAGAGAG 1537
Qy 501 CysArgLeuThrThrGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 520
Db 1538 TGTCAGCTGGCAACCCAGAGGCAAGAAATGAGAGCAACAAACAGCTTAGAGAGCTGAG 1597
Qy 521 IleAlaGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 540
Db 1598 ATTGGGAATACCCACTTACAGCAACAGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1657
Qy 541 LeuLysProGluLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 560
Db 1658 CTTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1717
Qy 561 HisArgAspSerLeuValThrLysLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1718 CATAGAGATTCGCTTCTTACCCTCAAAAGAGAGCTTGAAGCAAGAGAGAGAGAGAG 1777
Qy 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLysLysLys 600
Db 1778 CAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1837
Qy 601 AspilePheAsnAsnGlnLysGluLeuArgGluLysLysLysLysLysLysLysLys 620
Db 1838 GATGTTTCAACCAACAGCTGAAGGAACTGAGAGAGATACACAGTAAACAGAGCTCCAG 1897
Qy 621 LysGlnLysSerMetGluAlaGluArgLysLysLysLysLysLysLysLysLysLysLys 640
Db 1898 AAGCAGAGGTCATCGAGCGGAGAGAGCTTAAAGCAGAGAGAGAGAGAGAGAGAGAG 1957
Qy 641 GluLeuGluLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrp 660
Db 1958 GAGTTGGAGAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2017
Qy 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 2018 CAGGAGAGATGTCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2074
Qy 681 LysLeuLysArgGluGluSerValLysLysLysLysLysLysLysLysLysLysLysLys 700
Db 2075 AAACCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2134
Qy 701 AlaGlnAspLysLeuGluArgLysLysLysLysLysLysLysLysLysLysLysLys 720
Db 2135 GTGCAAGACAGCAGAGCTGGCTTTTCCATCCATCAAGAGAGAGAGAGAGAGAGAGAG 2191
```

```
Qy 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrLysSerAlaGlnGluAsnVal 740
Db 2192 CAGGACCCCTGGCCACAGCAGAGAGAGTCCGCTTACAACTCTCTGCACAGAGAGTGCC 2251
Qy 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluLeuThr 760
Db 2252 AAAGTGGTGTATTACCGAGCGCTGTACCCCTTTGAGTCCAGAGAGTCATGACGAGATCACC 2311
Qy 761 IleGlnProGlyAspIleValMet-----ValAspGluSerGlnThrGly 775
Db 2312 ATCCAGCCAGAGATATAGTCTAAAGGGGAGATGGGTGATGAAAGCCAGACCCGGA 2371
Qy 776 GluProGlyTyrLeuGlyGlyLysLysGlyLysThrGlyTyrPheProAlaAsnTyr 795
Db 2372 GAGCCAGGATGCTTTGGAGGAGAACCGAAGGGAGAGACAGAGATGTTCCCTTCGCAACATAT 2431
Qy 796 AlaGluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThr 815
Db 2432 GCAGAGAAAGATTCCAGAAATGAGATTCCCACTCCAGCCAAACCAAGTACCGATCTGACA 2491
Qy 816 SerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSer 835
Db 2492 TCTGCCCTTGGCCCAAACTGGCTCTGGGTGAGACCCCTGCTCTCTTCCAGTGACCTCT 2551
Qy 836 SerGluProSerThrThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSer 855
Db 2552 TCTGAGCCCTCCACAAACCCCAACAACTGGGAGAGTTCAGTTCCACGCTGGCCAGCAGC 2611
Qy 856 ThrAsnGluLysProGluThrAspAsnTrpAspAlaTrpAlaGlnProSerLeuThr 875
Db 2612 ACAAATGAGAAACAGAAACAGCAACTGGGAGACAGTGGGGGGCTCAGCCCTTCTCTGACT 2671
Qy 876 ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly 895
Db 2672 GTACCACTGTCTGGCCAGTTAAGCAGAGAGTTCGGCTTTCACACAGCCACAGCCACTGGC 2731
Qy 896 SerSerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAla 915
Db 2732 TCTTCCCTTCTCCGCTCTCGGCGCAGGGTGAAGAGTGAAGGGCTCCAAGCACAAGCC 2791
Qy 916 LeuTyrProTrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIle 935
Db 2792 CTGTATCTCTGGAGAGCCAAAGAGAACCACTTAAATTTTAAAGAGATGATGTATC 2851
Qy 936 ThrValLeuGluGlnGlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrp 955
Db 2852 ACCGTTCCTGGAAACAGCAGACATGTGTGTGTGGAGAGTTCAGAGTTCAGAGAGGTGG 2911
Qy 956 PheProLysSerTyrValLysLeuLysSerGlyProLysArgLysSerThrSerMetAsp 975
Db 2912 TTCCCAAGTCTTACGTGAAACTCATTTTCAGGGCCGCTAAGGAAATCCACAGCATCGAT 2971
Qy 976 SerGlySerSerGluSerProAlaSerLeuLysArgValAlaSerProAlaAlaLysPro 995
Db 2972 ACTGCGCTTACGAGAGCTCCCTCTAGTCTAAGAGAGTGGCTTCCCGAGCCGCGCAGCCA 3031
Qy 996 ValValSerGlyGluGluIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGlu 1015
Db 3032 GCCATTCTCGAGAGAAATCGCCAGGTTCATTGCTTCTCTACACTGTACGGGTCTCGAA 3091
Qy 1016 GlnLeuThrLeuAlaProGlyGlnLeuLysLeuLysArgLysLysAsnProGlyTrp 1035
Db 3092 CAGCTCACCTGGCTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3151
Qy 1036 TrpGluGlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTrpPheProAlaAsn 1055
Db 3152 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3211
Qy 1056 TyrValLysLeuLeuSerProGlyThrSerLysIleThrProThrGluProProLysSer 1075
Db 3212 TATGTCAAACTTCTAAGCCCTGGAACAGCAAAATCACCCCACTGAGTACCCAGAGCC 3271
Qy 1076 ThrAlaLeuAlaAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAsp 1095
```

```

Db      3272 GCAGTGCAGCCAGCAGTGTGCCAGGTGATCGGGATGTACGATACACCGCGCAGATGAC 3331
QY      1096 AspGluLeuAlaPheAsnLysGlyGlnIleAsnValLeuAsnLysGluAspProAsp 1115
Db      3332 GATGAGTACCTTTCAGCAAGCCAGATCATCAAGTCTCTAGCAAGAGGACCGGAC 3391
QY      1116 TrpTtpLysGlyGluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 1135
Db      3392 TGTGTGAAAGGAGAGTCACTGCGGCGAGGTGGGCTCTCTCCGTCCTCAATATTGTGAAGCTG 3451
QY      1136 ThrThrAspMetAspProSerGln 1143
Db      3452 ACCACAGCATGGACCCCGCCAG 3475

RESULT 9
AF132478
LOCUS   AF132478 3723 bp mRNA linear ROD 09-MAR-1999
DEFINITION Mus musculus Eael protein mRNA, complete cds.
ACCESSION AF132478
VERSION AF132478.1 GI:4378884
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3723)
AUTHORS Sengar,A.S., Wang,W., Bishay,J., Cohen,S. and Egan,S.E.
TITLE The EH and SH3 domain Eae proteins regulate endocytosis by linking to dynamin and Eps15
JOURNAL EMBO J. 18 (5), 1159-1171 (1999)
MEDLINE 99164083
PUBMED 10064583
REFERENCE 2 (bases 1 to 3723)
AUTHORS Sengar,A.S., Wang,W., Cohen,S., Bishay,J. and Egan,S.E.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1999) Programs in Cancer & Blood Research/Developmental Biology, The Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G-1X8, Canada

FEATURES
source
1..3723
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..3642
/codon_start=1
/product="Eael protein"
/protein_id="AADI9746.1"
/db_xref="GI:4378885"
/translation="MAQFPFPFGSLDVAITVEERAKHDQQLSLKPTAGFITGDOA
RNFQSGLPQVLAQIWAALADNMDQVDFEFSIAMKLIKLGQYQLPSTLPVPM
KQPVVAISSAPAFGIGIASMPLTAVAPVPMGSIPIVGMSPPLVSVPPAAVPLAN
GAPVTPQLPAPAPAAATWPKSSFSRSGPSOLNQLKQAKQFVDSAPPAEAWVP
QSSRLKROLFNHDKTMSGLTGPQARTILMOSSLPQQAQLASLWNLSIDIDDKLTA
EPIFLAMHLIDVAMSQPLPVPPIYIPPSFRVSGSGMSVSSSSVDQLRPEPS
SEDEQPEKKLPVTFEDKKRENFERSVELEKRRQALLQKQKBERLAQLERAQEER
KEREQEQAKRLQLEKQKRELERQEREERKEIEREAKRELERQLEWER
NRRQELINRQEQEVTVLKARKTLEFELEALNDKHQLEGLQDIDIRCLATQRE
IESTNKSRLRIAEITHLOQLQESQOQLRLPEKOILSDQLKQVQNSLRHDSLT
LKLALAKELARQLREQLDEVERTRSKLOEIDVFNOLKEIREIHSKQLOKORSL
EAARLKQEKERKSLELQKEDARQVRQDKWLEHVQEQEQPRPRPHEEDRLKR
EDSVRKEAEERAKMQDKSLRFLPHQEPKLAQAPWSTTEKGPLTISAQESVKV
VYRALYPFESSHSDITTPQGDIVMDESQTGPEGLGELKGTGFWFPFANYAKIP
ENEPPTKFPVPTDTSAPAKLALRETAPLPVTSEPTTNNMADFSSTWPSNSE
KPTDNMDTWAQPSLTVTSAGOLRQSAFTPATATGSSPSVLGQKEVQGLQAOAL
YPRAKNDHNLNFKNSDVLTVLQDDMMWFGEVQKQWFPKSVKLIISQVPKFSI
DGTPTESPLSKRVASPAKPAIPGEFTAMTYESESQGLDFQGGDVIIVTKDGD
WMTGTGKSGVFPFNNYKQSESGTAGTKSLGKPKPEIAQVIAASYAATGEQUTL
APQGLILIRKKRQIGELQARKKRIQFWFPANYVKLLSEPTSIPTPELPKTVL
QPAVCCQVIGWIDYTAQNDDDELAFSGQIINVLNKDPPDWKGEVSGVGLFSPNVKL
TTDMDFPSQ"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 6,53e-213 Length: 3723
Score: 5437.50 Matches: 1059
Percent Similarity: 90.28% Conservative: 37
Best Local Similarity: 87.23% Mismatches: 45
Query Match: 92.24% Indels: 73
DB: 10 Gaps: 3

US-09-720-934-2 (1-1143) x AF132478 (1-3723)

QY      1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db      1 ATGGCTCAGTTTCCACACCTTTCGGTGGTAGCTGGATGCTCGGCCATACTGTGGAG 60
QY      21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db      61 GAAAGGGCCAAAGCATCACCAGCAGTTCCTTAGCTGAGCCGATAGGGGATTTATACT 120
QY      41 GlyAspGlnAlaAArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db      121 GGTGATCAAGCGAGGAACTTTTTCCTCAATCTGGGTACCTCAGCTGTCTTAGACAA 180
QY      61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db      181 ATATGGCGCTAGCGACATGAATAACGATGAAGGATGATCAAGTGAATTTCCATA 240
QY      81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db      241 GCATGAAGCTTATCAAACTGAAGCTACAAGGATATCAGTCCCTCCACACTTCCCT 300
QY      101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db      301 GTCATGAAACAGCAACAGCTGGCTATTTCCAGTGCACCAAGCATTTGGTATAGAGG 360
QY      121 AlaSerMetProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db      361 GCTAGCATGCCACCATCAGCTGTGTCTCTGTGCTCAATGGCTCCCATTCAGTGTGT 420
QY      141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db      421 GGAATGTCTCCACCTTAGTATCTTCTGCTCCCTCCAGCAGCAGTGCCTCCCTGGCTAAC 480
QY      161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db      481 GGCGCTCTCCCGCTCATACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY      181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db      541 AGAGTTCCTCTTCCAGCAGATCTGCTCCAGGTCACAAATAAACACTAAGTTACAGAG 600
QY      201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
Db      601 GCACAATCATTCGATCTGCCAGCGCCCTCCAGCAGCAGAGATGGCTGTGCTCAGTCA 660
QY      221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db      661 TCAAGCTCAAAATACAGCAGGATTTATTCACAGCAGCAGCAAAAATATGATGAGTGA 720
QY      241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
Db      721 ACAGTCCCCAGCAGAACTTATTCATGCAATCAAGTTTACCCAGGCTCAGCTGGCT 780
QY      261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
Db      781 TCAATATGAAATCTTCTCAGATTCATGAATGGAATACTCACTGCAGAGAAATTAATC 840
QY      281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
Db      841 CTAGCTATGACCTTATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY      301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
Db      901 CCAGAAATACCTCCTCTCTTCCAGAGAGTTCGCTCCGCGAGTGGGATGTCCTGCATATA 960

```


QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
DB 961 AGCTCTTCTCTGTGGATCAGAGGCTGCTGAGAGCGCTGCTGACAGAGATGAGCAG 1020
QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
DB 1021 CCA--GAGAGAAACTGCTGTGACATTTGAAGATAAGAGCGGGAGAACTTCGAGCGA 1077
QY 361 GlyAsnLeuGluLeuGluLysArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
DB 1078 GGCAGTGTGAGCTGAGAGAGCGCGCAAGCGCTCTTGGAGCAGCAGCGCAAGAGCAG 1137
QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGlnGlu 400
DB 1138 GAGCGGTGGCTCAGCTGGAGCGCGCGAGCAGAGAGAAAGCGGGAGCGCCAGGAG 1197
QY 401 GlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
DB 1198 CAGAGGCCAAGCGGACGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAG 1257
QY 421 ArgGlnArgGluGluGluArgGlyLysGluLeuArgGluAlaAlaLysArgGlu 440
DB 1258 CGGCAGCGCAGAGAGAGAGAGAGAGATCGAGAGGCGCGAGGCGCGCAAAACGGGAA 1317
QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgGlnGluLeuAsnGln 460
DB 1318 CTGAAAGGCGAGCGACACTTGAATGGGAACCGGAACCGGAGACAGGAACCTCTGATCAG 1377
QY 461 ArgAsnLysGluGlnGluAspIleValLeuLysAlaLysLysLysThrLeuGluPhe 480
DB 1378 AGGNACAGAGAGCAGGAGGCGCACCTGGTCTCTGAAGCGCAGGAGAGACTCTGGAGTTT 1437
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
DB 1438 GAGTTAGAGCTCTGAATGACAAAGCATCAGCTAGAGGAAACCTTCAGGATATCAGG 1497
QY 501 CysArgLeuThrGlnArgGlnGluLeuLeuSerThrLeuLysSerArgGluLeuArg 520
DB 1498 TGTGACTGGCAACCCAGAGCGCAGAAATTGAGAGCAGCAACAAGTCTAGAGACTAAGA 1557
QY 521 IleAlaGluLeuThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnMetLeuGlyArg 540
DB 1558 ATTGCTGAATCACCCACTTACAGCAGCAGTGTGAGGAATCTCAGCAAAATGCTTGAAGA 1617
QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
DB 1618 CTTATTCCAGAGAAACAGATCTCAGTGACCAGTTAAACAAAGTCCAGCAGAACAGTTTG 1677
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
DB 1678 CATAGAGACTCGCTCTTACCCTCAAAAGACCTTGGAGCAAGAGAGCTGGCCCGGAG 1737
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
DB 1738 CAGCTCCGGGAGCAGCTGGAGCGGTGGAGAGAGAGACAGGCTCAAGCTCAGAGATT 1797
QY 601 AspIlePheAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
DB 1798 GATGTTTCAACACCGACTGAAGAACTGAGAGATACATAGCAACAGCAACTCCAG 1857
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIle 640
DB 1858 AAGCAGAGTCTCTGGAGCAGCGGCTGAGCAGAAAGAGAGCAGGAGAGAGAGCGCTG 1917
QY 641 GluLeuGluLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrp 660
DB 1918 GAGTTAGAGAAACAAAGAGACGCTCAGAGACGAGTTTCAGGAAAGGAGCAAGCAATGG 1977
QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGlu 680
DB 1978 CTGAGAGATGTGCAGCAGGAG--GAGCAGCCAGCCCGGAAACCCCAAGAGAGGAC 2034

QY 681 LysLeuLysArgGluGluSerValLysLysLeuAspGlyGluGluLysGlyLysGlnGlu 700
DB 2035 AGACTGAGAGGAGAGACAGTGTGAGGAAGAGCGCGGAGAGAGCAAGCGGAA 2094
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnGlnGluProAlaLysProAlaVal 720
DB 2095 ATCAAGACAAAGCAGAGTGGCTTTCCATCCCATCAGAGCCAGCTAAGCTGGCCACC 2154
QY 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
DB 2155 CAGGACCCCTGGTCTACACAGAGAAAGGCCCTTACCAATTTCTGCAGAGGAGGTGA 2214
QY 741 LysValValTrpTrpArgAlaLeuTrpPheGluSerArgSerHisAspGluIleThr 760
DB 2215 AAAGTGTATATTACCGAGCGCTGTACCCCTTTGAATCCAGAAAGTACAGATGATCACC 2274
QY 761 IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeu 780
DB 2275 ATCCAGCCAGAGATATAGTCAATGGATGAAGCCAGACTGGAGAGCCAGGATGGCTT 2334
QY 781 GlyGlyGluLysLysGlyLysThrGlyTrpPheProAlaAsnTrpAlaGluLysIlePro 800
DB 2335 GAGAGAGCTGAAGAGGAAAGACGGATGGTTCCTGCAAACTATGCAGAAAGATTCCA 2394
QY 801 GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro 820
DB 2395 GAAATGAGGTTCCCACTCCAGCCAAACAGTACCGATCTGACATCTGCCCTGCCCCC 2454
QY 821 LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr 840
DB 2455 AAATCGCTCTGGCTGAGACCCCTGCTCTTGCAGTGAGCTCTCTGAGCCCTCCACA 2514
QY 841 ThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysPro 860
DB 2515 ACCCCCAACACTGGCGAGACTTCAGTTCACGTGGCCCGCAGCAGCTCAAGCAGAGCCA 2574
QY 861 GluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThrValProSerAlaGly 880
DB 2575 GAAACGCAACATGGGATACGTGGCGGCTCAGCCTTCTCTGACCGTACCTAGTGTGGC 2634
QY 881 GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrArgLysSerSerPro 900
DB 2635 CAGTTACGGCAGAGATCAGCTTTTACCCCGCAGCCACAGCTGGCTCTCTCCCATCTCCC 2694
QY 901 ValLeuGlyGlnGlyGluLysValGluGlyLeuGlnAlaGlnAlaLeuTrpTrpArg 920
DB 2695 GTCCTGGCCAGGCTGAAGAGTGAAGGCTACAGCGCAGCCCTGTATCCCTGGAGA 2754
QY 921 AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGln 940
DB 2755 GCCAAAAAGACAAACCACTTAAATTTTAAACAAAAGTACGCTCATCCGTTCTGGAAACAG 2814
QY 941 GlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrpPheProLysSerTrp 960
DB 2815 CAAGACATGTGTGGTGTGGAGAGTTCAGGTCAAGGTCAAGAGGGTTGGTTCCTCCCACTCTAC 2874
QY 961 ValLysLeuIleSerGlyProIleArgLysSerThrSerMetAspSerGlySerSerGlu 980
DB 2875 GTGAACTCATTTCCAGGCCCGTGAAGAAATCCACAAAGCATCGATCTAGTGGCCCTACTGAA 2934
QY 981 SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu 1000
DB 2935 AGTCTGTAGTCTAAAGAGAGTGGCTTCCCGCGCCGCAAGCCAGCCATTCCTCCCGGAA 2994
QY 1000 ----- 1000
DB 2995 GAGTTTATGCCATGTACATACGAGGTTCTGAGCAGAGAGATTAAACCTTTACGAA 3054
QY 1000 ----- 1000
DB 3055 GGGGATGTGATTGTGGTTTACCAAGAAAGATGCTGCTGGTGAACGGGAAACGCTGGCGCAG 3114
QY 1000 ----- 1000

Db 3115 AAGTCGGAGCTTCCCTCTTAATATGTGAGGCTTAAAGATTACAGAGGCTCTGGAAC 3174
 QY 1001 -----GlulleAlaGlnValIleAlaSerTyr 1009
 Db 3175 GCTGGGAAAACAGGAGCTTTAGGAAAAAACCTGAATTCGCCAGGTTATGCTTCTTAC 3234
 QY 1010 ThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuIleLeuArgLys 1029
 Db 3235 GCTGCTACTGCTCCGAAACACTCACCTCGCTCTCGGAGCTGATCTGATCCGAAA 3294
 QY 1030 LysAsnProGlyGlyTyrTrpGluGlyGluLeuGlnAlaArgGlyLysArgGlnIle 1049
 Db 3295 AAGAACCCAGTGATGCTGAGAGGAGAACTGCAAGCTCGAGGAAAAAGCGCCAGATA 3354
 QY 1050 GlyTrpPheProAlaAsnTyrValLysLeuSerProGlyThrSerLysIleThrPro 1069
 Db 3355 GGGTGGTTTCCAGCAATTAATGTCAAACTTCTAAGCCCGGACACAGCAAAATCACCCCA 3414
 QY 1070 ThrGluProProlysSerThrAlaLeuAlaValCysGlnValIleGlyMetTyrAsp 1089
 Db 3415 ACTGAGCTACCCAGACCCGAGTCGACCCAGCAGTGTGCGAGGTGATCGGAGTACGAT 3474
 QY 1090 TyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLysGlyGlnIleIleAsnValLeu 1109
 Db 3475 TACACGCCCAAGACGATGACGACTAGCCTTCAGCAAGAGCCAGATCATCAACGTCCTC 3534
 QY 1110 AsnLysGluAspProAspTyrTrpLysGlyGluValAsnGlyGlnValGlyLeuPhePro 1129
 Db 3535 AACAGGAGGACCCGAGCTGTGGAAAGAGAGAGTCACTGGGCAAGTGTGGCTCTTCCCA 3594
 QY 1130 SerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
 Db 3595 TCCAATTATGAAAGCTGACACAGACATGAGACCCAGCCAGCCAG 3636

RESULT 10

AF132481
 LOCUS AF132481 5145 bp mRNA linear ROD 09-MAR-1999
 DEFINITION Mus musculus Ecell protein mRNA, complete cds.
 ACCESSION AF132481
 VERSION AF132481.1 GI:4378890
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 SENDER A.S., Wang, W., Bishay, J., Cohen, S. and Egan, S.E.
 TITLE The EH and SH3 domain Ese proteins regulate endocytosis by linking
 to dynamin and Eps15
 JOURNAL EMBO J. 18 (5), 1159-1171 (1999)
 MEDLINE 99164083
 PUBMED 10064583
 REFERENCE 2 (bases 1 to 5145)
 SENDER A.S., Wang, W., Cohen, S., Bishay, J. and Egan, S.E.
 DIRECT SUBMISSION
 TITLE Submitted (02-MAR-1999) Programs in Cancer & Blood
 Research/Developmental Biology, The Hospital for Sick Children, 555
 University Avenue, Toronto, ON M5G 1X6, Canada
 LOCATION/Qualifiers
 1. .5145
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1. .5145
 /codon_start=1
 /product="Ecell protein"
 /protein_id="AAD19749.1"
 /db_xref="GI:4378891"
 /translation="MAQFTPTPFPGSLVDVVAITVEERAKHDQQLSLKPIAGFTIGDQA
 RNFQFQSLQPPVLAQIWMALADNDRMDQVEFSIAMKLIKLGQYGLPSTLPPVM
 KOQVVAISSAPAFGIGIASMPPLTAVAPVMGSIPIVGMSPPLVSSVPPAAVPLAN
 GAPPVQLPAPAHPAATWPKSSSFSSRSGSLNTKLQKASFDVASAPPAEAWNP

FEATURES

source

CDS

ORIGIN

Alignment Scores:
 Pred. No.: 9,47e-213 Length: 5145
 Score: 5437.50 Matches: 1059
 Percent Similarity: 90.28% Conservative: 37
 Best Local Similarity: 87.23% Mismatches: 45
 Query Match: 92.24% Indels: 73
 DB: 10 Gaps: 3

US-09-720-934-2 (1-1143) x AF132481 (1-5145)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
 Db 1 ATGGCTCAGTTTCCACACTTTCGGTGTGAGCTCGATGCTCGGCGCATTAATGTGGAG 60
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 61 GAAAGGCGCAAGCATCACGACAGTTCCTTAGCCTGAAGCCGATAGCGGATTTATTACT 120
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 121 GGTGATCAAGCGAGGAACTTTTTCCTCAATCTGGGTACCTCAGCTGTCTTAGGCACAA 180
 QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 181 ATATGGCGCTAGCGACATGATACGATGAGGATGATGATGATGATGATGATGATGATGAT 240
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 241 GCCATGAAGCTTATCAAACTGAAGCTACAGGATATACAGTCCCTCCACACTTCCCTCT 300
 QY 101 ValMetLysGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 Db 301 GTCATGAAACAGCAACAGTGGCTATTTCCAGTGCCACGACATTTGGTATAGAGGAGATT 360
 QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 361 GCTAGCATGCCACCATCACACTGTGTCTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 420
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 Db 421 GGAATGTCTCCACCTTAGTATCTTCTGTCTCTCTCCAGCAGCAGTGTCTCTCCCTGCTAAC 480
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 Db 481 GGGGCTCTCTCCCTGTCATACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY	181	LySerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys	200
DB	541	AAGAGTCTTCTCCCTTCAGCAGATCTGCTCAGGGTCACAATTAACACTAAGTTACAGAAAG	600
QY	201	AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer	220
DB	601	GCACATCATTCGATGTCCAGCGCCCTCCAGCAGCAGATGGCTGTGCCTCAGTCA	660
QY	221	SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu	240
DB	661	TCAAGGCTGAAATACAGGCAGTTATTCAACAGCCACGACAAACATATCAGTGGACACTTA	720
QY	241	ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla	260
DB	721	ACAGGTCCTCCAGGCAAGAACTATTCTCATGCAATCAAGTTTACCCAGGCTCAGTGGCT	780
QY	261	SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGlnLeuPheIle	280
DB	781	TCAATATGGAATCTTTCTGACATTGATCAAGATGGAAAACTCACTGCAGAGAATTTATC	840
QY	281	LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProIleProValLeuPro	300
DB	841	CTAGCTATGCACCTAATTGATGTTCGATGTCTGGTCAGCCACTGCCGCCGCTCTCTGCT	900
QY	301	ProGluTrpIleProProSerPheArgValArgSerGlySerGlyIleSerValIle	320
DB	901	CCAGATATCATCTCTCTCTTCAGAGAGTTCCGCTCCGGCAGTGGGATGCCGTCATA	960
QY	321	SerSerThrSerValAspGlnArgLeuProGluGluProValLeuLeuAspGluGlnGln	340
DB	961	AGCTCTTCTTCTGTGGATCAGAGCTGCTTCGAGAGCCGCTGTCAGAGATCAGCAGCAG	1020
QY	341	GlnLeuGlnLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg	360
DB	1021	CCA--GAGAAAGAACTCCCTGTGCATTTTGAAGATAAGAACGCGAGAACTTCGAGCGA	1077
QY	361	GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuGluGlnGlnArgLysGluGln	380
DB	1078	GGCAGTGTGGAGCTGGAGAGCGCGGCAAGCGCTCTTGGAGCAGCAGCGCAAGAGCAG	1137
QY	391	GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu	400
DB	1138	GAGCGTTGGCTCAGCTGGAGCGCGCGCAGCAGAGAGAAAGCGGAGCGCCAGGAG	1197
QY	401	GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu	420
DB	1198	CAGAGGCCAAGCGCGAGCTGGAGTGGAGACGAGCTGGAGAACGAGCGCGGAGCTGGAG	1257
QY	421	ArgGlnArgGluGluArgArgLysGluIleGluArgGluAlaLysArgGlu	440
DB	1258	CGGCAGCGCAGAGGAGGAGAGAGAGAGATCGAGAGCGCGAGCGCCGAAACCGGAA	1317
QY	441	LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGluLeuLeuAsnGln	460
DB	1318	CTGGAAAGCGCAGCAGCACTTGAATGGAAACGGAACCGGAGACAGGAACCTCTGTAATCAG	1377
QY	461	ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe	480
DB	1378	AGGAACAAGGAGCAGGAGCGGCACCGTGTCTCTGAAGCAAGAGGAGAGACTCTGGAGTTT	1437
QY	481	GluLeuGluAlaLeuAsnAspLysIleHisGlnLeuGluGlyLysLeuGlnAspIleArg	500
DB	1438	GAGTTAGAGCTCTCTGATATGACAAAGCATCAGCTAGAGGAAAACTTCAGGATATCAGG	1497
QY	501	CysArgLeuThrThrGlnArgGlnIleGluSerThrAsnLysSerArgGluLeuArg	520
DB	1498	TGTCAGCTGGCAACCCAGAGGCAGAAATTCAGAGCGCAAGCAACAGCTTAGAGAGCTAAGA	1557
QY	521	IleAlaGluIleThrHisLeuGlnGlnLeuGlnLeuSerGlnGlnMetLeuGlyArg	540
DB	1558	ATTGCTGAATCACCACCTTACAGCGCATGTTCAGGAAATCTCAGCAANAATCTTGGAGA	1617

QY	541	LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu	560
Db	1618	CITATTCCAGAGAAAACAGATCTCACTGACCAAGTTAAACAACAGTCCAGCAGAACAGTTTG	1677
QY	561	HisArgAspSerLeuValThrIleuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln	580
Db	1678	CATAGAGACTCGCTTCTTACCTCAAAGAGCCTTGGAGCAAGAGAGCTGGCCCCGGCAG	1737
QY	581	HisLeuArgAspGlnLeuAspGluValGlnLysGluThrArgSerLysLeuGlnGluIle	600
Db	1738	CAGCTCCGGGAGCAGCTGGACGAGTGGAGAGAGACACAGGTCAAAGCTCAGGAGATT	1797
QY	601	AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln	620
Db	1798	GATGTTTCAACAACAGCTGAAGAACTGAGAGAGATACATAGCAACAACAGCAACTCCAG	1857
QY	621	LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnArgLysIleIle	640
Db	1858	AAGCAGAGGTCCCTGGAGGACGCGCACTGAAGCAGAAAGCAGGAGAGAGAGAGCCTG	1917
QY	641	GluLeuGluLysGlnLysGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrp	660
Db	1918	GAGTTAGAGAGCAAAAGAAAGACGCTCAGACAGCAGTTCAGGAAAGGACAGCAATGG	1977
QY	661	LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu	680
Db	1978	CTGGAGCATGTGCAGCAGGAG--GAGCAGCCACGCCCGCCGAAACCCACGAGAGAGC	2034
QY	681	LysLeuLysArgGluGluSerValLysLysIleAspGlyGluGluLysGlyLysGlnGlu	700
Db	2035	AGACTGAAGAGGGAAGACAGTGTTCAGAAAGAGAGGCGGAAGAGAGAGCCCAAGCCGGAA	2094
QY	701	AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal	720
Db	2095	ATGCAAGACAGCAGAGATCGGCTTTTCATCCGCATCAGAGCCAGCTAAGTCGCCACC	2154
QY	721	GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal	740
Db	2155	CAGCACCTGTCTACACAGAGAAAGGCCGCTTACCATTTCGTGCACAGAGAGTGTA	2214
QY	741	LysValValTyrrArgAlaLeuTyrrProPheGluSerArgSerHisAspGluIleThr	760
Db	2215	AAAGTGGTATTATCCAGAGCGCTGTACCCCTTTGAATCCAGAGTCACGATGAGATCACC	2274
QY	761	IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeu	780
Db	2275	ATCCAGCCAGAGAGATATAGTCATGCTGGATGAAGCCAGACTGGAGGCCAGATGGCTT	2334
QY	781	GlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIlePro	800
Db	2335	GGAGGAGAGCTGAAAGGGAACGCGATGTTCCCTGCAAACTATGCAGAAAAGATTCCA	2394
QY	801	GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro	820
Db	2395	GAAATATGAGTTCCACATCCAGCCAAAACAGTGACCGCATCTGCACATCTGCCCTGCCCCC	2454
QY	821	LysIleAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr	840
Db	2455	AAACTGGCTCTGCGTGAGACCCCTGCTCTTTTGCAGCTGACCTCTTCTGAGCCCTCCACA	2514
QY	841	ThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysPro	860
Db	2515	ACCCCAACAACCTGGCGAGACTTCAGTTCCAGTGGCCGACGAGCTCAAAACAGAGAAGCA	2574
QY	861	GluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThrValProSerAlaGly	880
Db	2575	GAAACGGCAACTGGGATACGTGGCGCGCTCTCTGACCGCTACTAGTGTCTGCTGCTG	2634
QY	881	GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro	900
Db	2635	CAGTTACGGCAGAGATCAGCCTTTACCCCAACAGCCACTGGCTCTCTCCCATCTCCCT	2694
QY	901	ValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAlaLeuTyrrProTrpArg	920

QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 221 ATATGGCGCTGGCTCACATGAATAAGGACGGAAGGATGGATCAGGTGGAGTTTTCATA 280
QY 81 AlaMetLysLeuLeuLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 281 GCCATGAAGCTCATCAACTGAAGCTGCAAGGATATCAGTCCCGCCGACATCCCGCT 340
QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db 341 GTCATGAAGCAGCAGCAGCGCGCCATCTCTAGTGCACACAGCGTTTGGTATAGAGGGATG 400
QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db 401 GCTGGAAATGCAACCTGACAGCTGTGTCTCCCGTGCATATGGCTCCATCCCGATGTGT 460
QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db 461 GGAATGTCTCCGCCCTTAGTATCTTGTCTCCCTCAACGACAGTGGCTCCCTGGCTAAC 520
QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db 521 GGGGCCCTCTCTGTATACAGCTCTGCTGCTATTTGCTATCTCATCTGAGCCACATTCGCA 580
QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db 581 AAGAGTCTCTCTTACAGCAGATCTGCTCCAGGGTCACAAATTAACACTAAACTACAGAG 640
QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
Db 641 GCACAACTATTGATGTAGCAGCGCCCTGCAGCGGCAGAAATGGCTGTGCTCAGTGC 700
QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db 701 TCAAGACTGAATACAGCAGTATTCAACAGTCACGACAGACCAAGACCATGAGTGGACATTA 760
QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
Db 761 ACAGGTCCCCAGGCAAGAACTATTCTCATGCAATCAAGTTTACCCAGGCTCAGCTGGCT 820
QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
Db 821 TCAATCTGAATCTTCCGACATGATCAAGATGGAAGCTCACCGCAAGAAATTCATC 880
QY 281 LeuAlaMetHisLeuLeuAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
Db 881 CTAGCGATGCACCTGATTCATGTTGCTATGCTGCTCAGCCACTGCCCGCTGTCTGCT 940
QY 301 ProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle 320
Db 941 CCAGAATAATACCTCCCTTCTTTAGAAAGTTTCTGCTGAGGAGTGTGGATGTCCGTCTATA 1000
QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuLeuAspGluGlnGln 340
Db 1001 AGCTCTTCGTCTGCAGACCGCGCTGCCGAGGAGCCATCTCTCAGAGGATGAG---CAG 1057
QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
Db 1058 CAGGTGGAAAGAAAGCTGCTGTGACATTTGAAGATAAGAGCGGGAACCTTCAGGCGA 1117
QY 361 GlyAsnLeuLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
Db 1118 GGCAACCTGGAGCTGGAGAGCGCAGCGAGGCGCTCTCTGGAGCAGCGCAGAGGAGCAG 1177
QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu 400
Db 1178 GAGCGCTGGCTCAGCTGAGCGCGCGCAGCAGAGAGAGAGAGAGCGGAGCGCGCGAGAA 1237
QY 401 GlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGlnLeuGlu 420
Db 1238 CAGGAGCGCAAGACAGCTAGAGCTGGAGAGCAGCTAGAAAGCAGCGGAGCTGGAG 1297
QY 421 ArgGlnArgGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu 440

Db 1298 AGGCACAGAGAGAGGAGAGAGAGGATCTCAGAGCGCGGAGGAGCAAAACCGGAG 1357
QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgGlnGlnGluLeuLeuAsnGln 460
Db 1358 CTCGAACCGCAGCAGCAGCTTGAATGGAAACCGAACCGGAGACAGGAATCTCTGACTCAG 1417
QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
Db 1418 AGGAACAGAGACAGGAGGCGCATCTGTCTCTCAGGCGAGGAGGAAGACTCTGGAGTTT 1477
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
Db 1478 GAATTAGAGCTCTGAATGACAAAAGCATCAGCTGGAAGAAATCTTCAGGATATCAGG 1537
QY 501 CysArgLeuThrGlnArgGlnGluLeuIleGluSerThrAsnLysSerArgGluLeuArg 520
Db 1538 TGTGCACTGCAACCCAGAGCGCAAAATTTGAGAGCAAAACAAAGTCTTAGAGAGTGA 1597
QY 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGlnGlnMetLeuGlyArg 540
Db 1598 ATTGCGGAATCACCACATTACAGCAACAGTTGCAGGAATCTCAGCAGATGCTTGAAGA 1657
QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1658 CTTATTCCAGAGAGCAGATACTCAGTACCAGTTAAACAAAGTCCAGCAGAACAGCTTG 1717
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1718 CATAGAGATTGCGCTTCTTACCTCCAAAGAGCCTTGAAGCAAAAGAACTGGCCCGCAG 1777
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
Db 1778 CAGCTTCGAGAGCAGCTGGAGAGTGGAGAAAGAGACAGACAGTCAAAAGCTCAGAGATT 1837
QY 601 AspilePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
Db 1838 GATCTTTTCAACACACAGCTGAAGAACTTGAGAGAGATACACAGTAAACAGCAGCTCCAG 1897
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIle 640
Db 1898 AAGCAGAGCTCTCAGGCGCGAGAGCTTAAAGCAAGAGAGCAGGAGGAGAGAGCCTG 1957
QY 641 GluLeuGluLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrp 660
Db 1958 GAGTTGGAGAGCAAAAGAGAAAGGTCAGAGACAGTTCAGGAAAGGAGCAAGCAATGG 2017
QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 2018 CAGGAGCATGTCCAGCAGAG---GAGCAGCAGCGCCCTTCGGAACCCACAGAGAGGAC 2074
QY 681 LysLeuLysArgGluGluSerValLysLysAspGlyGluGluLysGlyLysGlnGlu 700
Db 2075 AAACCTGAAAGAGAGACAGTGTCAAGAAAGAGGCGGAGAGAGAGAGAGCCAGCGGAA 2134
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal 720
Db 2135 GTGCAAGCAAGCAGAGTCCGCTTTTCCATCCATCAAGAGCAGCAGTCAAGCGCGC--- 2191
QY 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
Db 2192 CAGGACCCCTGCCCCACAGCAGAGAAAGGTCGCTTCAATCTCTGCAGAGAGAGTGCC 2251
QY 741 LysValValTyrTrpArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
Db 2252 AAAGTGGTGTATTCAGGAGCGCTGTACCCCTTTGAGTCCAGAAAGTCATGACGAGATCACC 2311
QY 761 IleGlnProGlyAspIleValMet-----ValAspGluSerGlnThrGly 775
Db 2312 ATCCAGCCAGGAGATATAGTTCATGTTTAAAGGGGAATGGGTGATGAAAGCCAGACCGGA 2371
QY 776 GluProGlyTrpLeuGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyr 795

/gene="DKFZp686J17173"

4289

/gene="DKFZp686J17173"

polya_site

ORIGIN

Alignment Scores:

Pred. No.: 2,48e-197 Length: 4321
Score: 5057.00 Matches: 981
Percent Similarity: 99.90% Conservative: 0
Best Local Similarity: 99.90% Mismatches: 1
Query Match: 85.78% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2 (1-1143) x HSM806384 (1-4321)

QY 162 AlaProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLys 181
DB 51 GCTCCCTGTTATACAACTCTGCTGCAATTCCTGTCGACCCACATTCGCAAG 110
QY 182 SerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAla 201
DB 111 AGTTCTTCCTTTAGTAGATCTGCTCCAGGGTCACAACTAAACACTAAATTACAAAGCA 170
QY 202 GlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSerSer 221
DB 171 CAGTCAATTGATGTGGCCAGGTGCCACAGTGGCAGAGTGGCTGTTCTCAGTCATCA 230
QY 222 ArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThr 241
DB 231 AGACTGAATACAGCAATTTCAATAGTATGACAAACTATGAGTGGACACTTAACA 290
QY 242 GlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSer 261
DB 291 GGTCCCAAGCAAGAACTATTCTTATGAGTCAAGTTTACCACAGGCTCAGCTGGCTTCA 350
QY 262 IleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluPheIleLeu 281
DB 351 ATATGGAACTCTTCTGACATTTGATCAAGATGAAACTTACAGCAGAGAAATTTATCCTG 410
QY 282 AlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuProPro 301
DB 411 GCAATGACCTCATGTATGTAGTATGTCGGCAACCACTGGCCACTGTCGCTCCA 470
QY 302 GluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIleSer 321
DB 471 GAAATACATTCACCTCTTTTAGAAGAGTTCGATCTGGCAGTGGTATATCTGTATAAGC 530
QY 322 SerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 341
DB 531 TCAACATCTGTAGATCAGAGCTACAGAGAACCCAGTTTTAGAAGATGAACAAACAA 590
QY 342 LeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArgGly 361
DB 591 TTGAAAGAAATTAACCTGTAAACGTTTGAAGTAAAGACGGAGAACCTTTGAACGTGGC 650
QY 362 AsnLeuGluLeuLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGlnGlu 381
DB 651 AACCTGGAACCTGGAGAAACGAAAGCAAGCTCTCTCTGGAACAGCAGCGCAGCAGAG 710
QY 382 ArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGln 401
DB 711 CGCTGGCCACAGCTGGAGCGGGCGAGCAGAGAGGAGGAGCGTGGAGCGCAGAGCAA 770
QY 402 GluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuArg 421
DB 771 GAGCGCAAAAGACAACTGGAATCGAGAGAGCAACTGGAAAGACAGCGGAGCTGAACAG 830
QY 422 GlnArgGlnGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGluLeu 441
DB 831 CAGAGAGAGGAGGAGAGGAGAAAGAAATTCAGAGGGCGAGAGGCTGCAAAACGGGAAC 890
QY 442 GluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGluLeuLeuAsnGlnArg 461

DB 891 GAAAGCAACGACAACTTTCAGTGGCAACGGAATCGAAGCAAGAACTACTAAATCAAGA 950
QY 462 AsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPheGlu 481
DB 951 AACAAAGAACCAAGAGGACATAGTTGTACTGAAGCAAAAGAAAGAACTTTGGAATTTGAA 1010
QY 482 LeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArgCys 501
DB 1011 TTAGAGCTCTTAATGATATAAAGCATCACTAGAGGGAACCTTCAAGATATCAGATGT 1070
QY 502 ArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArgIle 521
DB 1071 CGATTGACCAACCAAGCAAGAAATTCAGAGCACAAAACAATCTAGAGAGTTGAGAATT 1130
QY 522 AlaGluIleThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArgLeu 541
DB 1131 GCCGAAATCACCCACTCTACAGCAACAATTCAGGAATCTCAGCAATCTTTGGAAGACT 1190
QY 542 IleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHis 561
DB 1191 ATTCAGAAACACAGATACTCAATGACCAATTTAAACAAGTTTACAGACAACTTTGCAC 1250
QY 562 ArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHis 581
DB 1251 AGAGATTCACTTGTACACTTAAAGAGCCTTAGAAGCAAAAGAACTAGCTCGGAGCAC 1310
QY 582 LeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluLeuAsp 601
DB 1311 CTAGCAGACCAACTGGAGTGAAGTGGAGAAAGAACTAGATCAAACTACAGAGAGTTGAT 1370
QY 602 IlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGlnLys 621
DB 1371 ATTTTCAATATCAGCTGAAGAACTAAGAGAAATACACATAGCAACCACTCCAGAG 1430
QY 622 GlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleLeuGlu 641
DB 1431 CAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAGAAACGAAAGATCATAGAA 1490
QY 642 LeuGluLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrpLeu 661
DB 1491 TTGAAAGAAACAAAGAAAGAGCCCAAGACGAGCTCAGGAAGAGGCAACAGCAGTGGCTG 1550
QY 662 GluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlnLys 681
DB 1551 GAGCATGTGCAGCAGGAGGAGCAGCATCAGAGACCAAGAACTCCACAGAGAGAGAAAA 1610
QY 682 LeuLysArgGluGluSerValLysLysAspGlyGluGluLysGlyLysGlnGluAla 701
DB 1611 CTGAAAAGGGAGGAGAGTGTCAAAAAGAGGATGGCAGGAGGAAAAAGGCAACAGGAAGCA 1670
QY 702 GlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaValGln 721
DB 1671 CAAGCAGAGCTGGGTGCGCTTTTCCATCAACACAGAACCCAGCTAGCCAGCTGTCCAG 1730
QY 722 AlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnValLys 741
DB 1731 GCACCTGGTCCACTGCAGAAAAAGGTCCTTACCATTTCTGCACAGGAAAAATGAAAA 1790
QY 742 ValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIle 761
DB 1791 GTGGTGTATTACCGGGCACTGTACCCCTTTGAATCCAGAGACCCATGATGAATCACTATC 1850
QY 762 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeuGly 781
DB 1851 CAGCAGAGAGACATAGTTCATGGTGGATGAAGCTAAACTGGAGAACCCGCGCTGGCTTGG 1910
QY 782 GlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIleProGlu 801
DB 1911 GGAGAAATTAAGAGGAAAGACAGGCTGTTCTCTGCAAACTATGACAGAGAAATCCAGAA 1970
QY 802 AsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaProLys 821
DB 1971 AATGAGTTCCCGCTCCAGTGAACACAGTACTGATTCACATCTGCCCTGCCCCCAAA 2030

misc_feature 3607..3786
/note="encodes SH3E domain"

ORIGIN

Alignment Scores: Pred. No.: 5,18e-185 Length: 4103
Score: 4754.00 Matches: 937
Percent Similarity: 83.40% Conservative: 78
Best Local Similarity: 76.99% Mismatches: 114
Query Match: 80.64% Indels: 88
DB: 5 Gaps: 10

US-09-720-934-2 (1-1143) x AF032118 (1-4103)

QY 1 MetAlaGlnPheProThrPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
DB 193 ATGGCTCAGTTTGGAACTCGTTGGGGTAATTTGGACATCTGGGCCATAAGCGTGGAG 252
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
DB 253 GAACGAGCTAAACATCACCAGCAGTTTCCATGGGCTCAAGCCAAACAGCTGGATATATACA 312
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
DB 313 GGTGACCAAGCTAGGAAATTTTCTTCTAGTCCGGCTCCCGCAGCTGTGTGGCACAG 372
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
DB 373 ATATGGCCCTCGCTCACATGAACATGATGGAGGATGGACCACTGGAGTTCTCCATA 432
QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuPro 100
DB 433 GCTATGAATTAATCAAGCTAAACACTACAAGTTTACCCTGTGCTCTATTCTCCCTCC 492
QY 101 ValMetLysGlnGlnProValAlaIleSerSerAla-----ProAlaPheGlyMetGly 118
DB 493 AATATGTTGAAGCAACCGTTCGCAATGCTGCTGCTGTGGGGGTTCGGGATGAGT 552
QY 119 GlyIleAlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIlePro 138
DB 553 GGAATCGTGGGCAATTCCTCCCTAGCAGCTGTAGCCCTGTGCCAATGCCATCCATCCA 612
QY 139 ValValGlyMetSerProThrLeuValSerSerValProThrAlaAlaValProLeu 158
DB 613 GTAGTAGGAATGTCT 666
QY 159 AlaAsnGlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThr 178
DB 667 TCTAACGGGGCTCTCGCAGTAATACAGTCTCATCTCTGTTTCTGCTCAC---TCTGCCACA 723
QY 179 LeuProLysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeu 198
DB 724 TTGCCAAAGAGTTCTTCAATCGGCGCTCAGTTGCGGGTCCCAGATAAACACATAACTT 783
QY 199 GlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValPro 218
DB 784 CAGAAAGCTAGTCTCTCGAATGTTCCAGCCCCCTTCTGTTGAATGGGCTGTGCCA 843
QY 219 GlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGly 238
DB 844 TCGTCTTCAAGATTGAAGTACAGACAAATGTTCAACAGCCAGGCAAAACCATGAGTGGG 903
QY 239 HisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGln 258
DB 904 AATTTACAGCTCTCTCAGAGTTCAGACATTGACCAAGATGGAAATCTTACCGCAAGCAG 963
QY 259 LeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGlu 278
DB 964 CTAGCTACATATGGAATCTTTTACAGATTGACCAAGATGGAAATCTTACCGCAAGCAG 1023
QY 279 PheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProVal 298
DB 1024 TTTATATTAGCTATGCACTTAATAGATGTGGCCATGTCTGGCCAGCCACTTCTCCCAATC 1083

QY 299 LeuProProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSer 318
DB 1084 CTGCTCCAGAGTATATTCCTCCCATCTTTTAGAAGAGTTCGATCAGGAGTGGGTATCC 1143
QY 319 ValIleSerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGlu 338
DB 1144 ATTATGAGCTCTGTGCTGTGTATCAGCGCTTCCGAGGAAACAGAA---GAAGAGGAG 1200
QY 339 GlnGlnGlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPhe 358
DB 1201 CCACAAACCGCGACAAAAGCTCCAGTCACATTTGAAGACAGAAAAGGGAGAACTTT 1260
QY 359 GluArgGlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGlnGlnArgLys 378
DB 1261 GAGCGAGGAAACCTAGAAATTTAGAAAAGAGGAGAGCAGCACTTTTAGAGCAGCAGCGCAA 1320
QY 379 GluGlnGluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArg 398
DB 1321 GAGCAGAGAGGTTGGCTCAGCTGGAAAGGCGCAGAGAGAGAAAGAACGAGAGCGG 1380
QY 399 GlnGluGlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGlu 418
DB 1381 CAAGATCAAGAGCGGAAGCGCAGCAGCAAGCTGGAAGAAACAGCTGGAGAAACAAACGAGAA 1440
QY 419 LeuGluArgGlnArgGluGluArgLysGluIleGluArgArgGluAlaAlaLys 438
DB 1441 CTGGAGCGGCAACGGGAGAGAGGAGCGCAGAAAGAAATCGAAAGGAGAGAGCGCGCAAAA 1500
QY 439 ArgGluLeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgGlnGlnLeuLeu 458
DB 1501 AGAGAACTAGAAAGCAACCGCAGTTGGAGTGGGAACGAAACAGAAAGGCAAGCACTCTT 1560
QY 459 AsnGlnArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeu 478
DB 1561 AATCAAGGAATAGAGAGCAAGAGGACATTTGTTCTTAAAGCAAGAGAGAACTCTTA 1620
QY 479 GluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAsp 498
DB 1621 GAGTTTGAGCTGGAAAGCTCTAAATGATAAAGACACCACTTGGAAAGGAAATCTCCAGGAT 1680
QY 499 IleArgCysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGlu 518
DB 1681 ATCCGGTGTCCCTCCTACTCTAACCGCAGCAAAATAGAAAGCACTTAAACAAATCCAGGAA 1740
QY 519 LeuArgIleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGlnGlnMetLeu 538
DB 1741 CTGAGGATCGCAGAGATCACACACTTGCAGCAGCAGCTTCAGGAATCTCAGCAGCTGCTT 1800
QY 539 GlyArgLeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsn 558
DB 1801 GGAATAATGATCTCTGAGAAACAGTCCCTTATCGCATCAATTAAGAGCAAGTTCAACAAAC 1860
QY 559 SerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAla 578
DB 1861 AGTTTGACAGAGATTCCTTCTTACACTAAAGAGGCAATTTGAAACCAAGAGGATTTGGA 1920
QY 579 ArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGln 598
DB 1921 CGCCAGCAGCTCCGAGATCAGTTAGATGAAGTAAAGAAAGAAACCGCAGCTTAAGCTTCA 1980
QY 599 GluIleAspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGln 618
DB 1981 GAAATCGAGTATTTAATACCAATTAAGAACTGAGAGAAATTTGTACAAACAGCAGCAG 2040
QY 619 LeuGlnLysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLys 638
DB 2041 TTCCAGAAACAGCAGGAGCTTTGAAACTGAGAGAGATCAAGCAGAAAGAGCTTGGAAACGAAA 2100
QY 639 IleIleLeuLeuLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLys 658
DB 2101 ACCAGCAGCTGTGATAGCTGAAGAGGAGGAGCAAAAGGCGGATGCTAGAGCAGGACAG 2160

FEATURES

Location/Qualifiers

source

1. .3231
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1,12e-127 Length: 3231
 Score: 3344.00 Matches: 640
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 56.73% Indels: 0
 DB: 6 Gaps: 0

US-09-720-934-2 (1-1143) x BD205037 (1-3231)

QY 504 ThrThrGlnArgGlnGlnLeuLeuSerThrAsnLysSerArgGluLeuArgLleAlaGlu 523
 Db 2 ACCACCCAAAGGCAAGAAATTGAGAGCACAAACAATCTAGAGAGTTGAGAAATTGCCGAA 61
 QY 524 IleThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArgLeuLeuPro 543
 Db 62 ATCACCCTCTACAGCAACAATTTACAGGAATCTCAGCAAAATGCTTGGAGACTTATTCCA 121
 QY 544 GluLysGlnLleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAsp 563
 Db 122 GAAAAACAGATATCTCAATGACCAATTAACAAAGTTTCAGAGAACAGTTGACACAGAT 181
 QY 564 SerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArg 583
 Db 182 TCATTTGTTACATTTAAAGAGCTTTAGAGAGCAAAAGAACTAGTCGGCAGCACCTACGA 241
 QY 584 AspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGlnLleAspLlePhe 603
 Db 242 GACCAACTGGATGAAGTGGAGAAAGAACTAGATCAAACTACAGGAGATTGATATTTTC 301
 QY 604 AsnAsnGlnLeuLysGluLeuArgGluLleHisAsnLysGlnGlnLeuGlnLysGlnLys 623
 Db 302 AATAATCAGCTGAAGGAACCTAAGAGAAATACAAATAGCAACACTCCAGAGCAAAAG 361
 QY 624 SerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysLleLeuGluLeuGlu 643
 Db 362 TCCATGGAGCTGAACGACTGAAACAGAAAGAACAGAACGAAAGATCATAGAAATTAGAA 421
 QY 644 LysGlnLysGluGluAlaGlnArgAlaGlnGluArgAspLysGlnThrLeuGluHis 663
 Db 422 AAACAAAAAGAAAGCCCAAGACGAGCTCAGGAAAGGGACAGCAGTGGTGGAGCAT 481
 QY 664 ValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGluLysLeuLys 683
 Db 482 GTGCAGCAGGAGGACGAGCATCAGAGACCAAGAAACTCCAGAGAGGAAACTGAAA 541
 QY 684 ArgGluGluSerValLysLysAspGlyGluLysGlyLysGlnGluAlaGlnAsp 703
 Db 542 AGGAGGAGAGTGTCAAAAAGAGAGATGGCGAGGAAAAAGCAAAACAGGAAGCAACAAG 601
 QY 704 LysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaValGlnAlaPro 723
 Db 602 AAGCTGGTGGCTTTTCCATCAACACCAAGAACAGCTAAGCCAGCTGTCAGGCAACC 661
 QY 724 TrpSerThrAlaGluLysGlyProLeuThrLleSerAlaGlnGluAsnValLysValVal 743
 Db 662 TGGTCCACTGCAGAAAAGGTCCACTTACATTTCTGCACAGGAAATGTAAAAGTGGTG 721
 QY 744 TyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluLleThrLleGlnPro 763
 Db 722 TATTACCGGCCTGTACCTTGTGAATCCAGAAAGCCATGATGAAATCACTATCCAGCCA 781
 QY 764 GlyAspLleValMetValAspGluSerGlnThrGlyGluProGlyThrLeuGlyGlu 783
 Db 782 GGAGACATAGTCATGGTGGATGAAAGCCAAACTGGAGAACCCCGCTGGCTTGGAGGAGAA 841

QY 784 LeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysLleProGluAsnGlu 803
 Db 842 TTAAAGGAAGACACAGAGGTGGTTCCCTGCAAACTATGCAGAGAAATCCAGAAAATGAG 901
 QY 804 ValProAlaProValLysProValThrAspSerThrSerAlaProAlaProLysLeuAla 823
 Db 902 GTTCCCGCTCCAGTGAACACAGTGAATCAACATCTGCCCTCTGCCCAAACTGGCC 961
 QY 824 LeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThrThrProAsn 843
 Db 962 TTGCGTGAGACCCGCCCTTTGGCAGTAACTCTTCAGAGCCCTCCAGACCCCTAAT 1021
 QY 844 AsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysProGluThrAsp 863
 Db 1022 AACTGGCGGACTTCAGTCCAGTGGCCACAGCACGAATGAGAAACAGAAACCGAT 1081
 QY 864 AsnTrpAspAlaTrpAlaAlaGlnProSerLeuThrValProSerAlaGlyGlnLeuArg 883
 Db 1082 AACTGGGATGCAATGGCGACGCCCTCTCTCAGCTTCCAGTCCCGGCCAGTTAAGG 1141
 QY 884 GlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerProValLeuGly 903
 Db 1142 CAGAGTTCGCTTTACTCCAGCACGGCCACTGGCTCTCCCGTCTCTGTGTAGGC 1201
 QY 904 GlnGlyGluLysValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrpArgAlaLysLys 923
 Db 1202 CAGGGTGAAGAGTGGAGGGCTTACAAGCTCAAGCTTATATCTTTGGAGAGCAAAAA 1261
 QY 924 AspAsnHisLeuAsnPheAsnLysAsnAspValLleThrValLeuGlnGlnAspMet 943
 Db 1262 GACAACCACTTAAATTTTAAACAAAAATGATGTATCATCCGCTCTCGAACAGCAGCATG 1321
 QY 944 TrpTrpPheGlyGluValGlnGlyGlnLysGlyTrpPheProLysSerTyrValLysLeu 963
 Db 1322 TGTGTGTTGGAGAGTTCAGGTTCAGAGGGTGGTTCCCAAGTCTTACGTGAATC 1381
 QY 964 IleSerGlyProLysArgLysSerThrSerMetAspSerGlySerSerGluSerProAla 983
 Db 1382 ATTTCAAGGGCCCAATAGGAAGTCTACAAGCATGGATTTCTGGTCTTTCAGAGAGTCTGCT 1441
 QY 984 SerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGluGluAla 1003
 Db 1442 AGTCTAAGCGAGTAGCTCTCCAGCAGCCCAAGCGTCTTTCCGGAGAGAAATTTGCC 1501
 QY 1004 GlnValIleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGln 1023
 Db 1502 CAGTTATTGCTTCATACACCGCCACCGCCCGGAGCAGCTCACTCTCGCCCTGGTCAG 1561
 QY 1024 LeuIleLeuLysArgLysLysAsnProGlyGlyTyrTrpTrpGluGlyGluLeuAlaArg 1043
 Db 1562 CTGATTTTGATCCGAAAAAGAACCCAGGTGGATGGTGGAGAGAGAGCTGCAAGCACGT 1621
 QY 1044 GlyLysLysArgGlnLleGlyTyrPheProAlaAsnTyrValLysLeuLeuSerProGly 1063
 Db 1622 GGGAAAAAGCGCCAGATAGGTGTTCCAGCTTAATTTATGTAAAGCTTCTAAGCCCTGGG 1681
 QY 1064 ThrSerLysLleThrProThrGluProProLysSerThrAlaLeuAlaValCysGln 1083
 Db 1682 ACGAGCAAAATCACTCCACAGAGCCACCTAAGTCAACAGCATTTAGCGGAGTGTGCCAG 1741
 QY 1084 ValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLysGly 1103
 Db 1742 GTGATTTGGATGACGACTACACCGCGCAGAAATGACGATGAGTGGCTTTCACAAAGGGC 1801
 QY 1104 GlnIleLeuValLeuAsnLysGluAspProAspTrpTrpLysGlyGluValAsnGly 1123
 Db 1802 CAGATCATCAAGCTCTCAACAGAGAGACCTGACTGGTGGAAAGAGAGAGTCAATGGA 1861
 QY 1124 GlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
 Db 1862 CAAGTGGGCTCTTCCCATCAATTTATGTGAAGCTGACCAAGACATGAGCCCAAGCCAG 1921

RESULT 15


```
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
Db 1576 GAAATTAGAGCTCTAAATGATAAAGCATCAACTAGAGGGAAACTTTCAGATATCAGA 1635
QY 501 CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg 520
Db 1636 TGTCCATTGACCAACCCCAAGGCAAGAAATTGAGACACAAACAATCTAGACAGTTGAGA 1695
QY 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGlnGlnMetLeuGlyArg 540
Db 1696 ATTGCCGAATCACCCTCTACAGCAACAATACAGGAATCTCAGCAAAATGCTTGGAGA 1755
QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1756 CTTATTCCAGAAAAACAGATACTCAATGACCAATTAAACAAAGTTCCAGCAGAACAGTTTG 1815
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1816 CACAGAGATTCACTTGTACCTTAAAGAGCCCTTAGAGCAAAAGAACTAGCTCGGCAG 1875
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
Db 1876 CACCTACGAGACCAACTGGATGAAGTGGAGAAAGAACTAGATCAAAACTACAGGAGATT 1935
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
Db 1936 GATATTTTCAATAATCAGCTGAAGGAACCTAAGAGAAATACACAATTAAGCAACTCCAG 1995
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIle 640
Db 1996 AAGCAAAAGTCCATGAGGCTTAACGACTGAAACAGAAAGCAAGAACGAAAGATCATA 2055
QY 641 GluLeuGluLysGlnLysGluGlu 648
Db 2056 GAATTAGAAAAAAGAAAAA 2079
```

Search completed: August 3, 2004, 07:17:20
Job time : 9667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 3, 2004, 04:38:58 ; Search time 4208 Seconds
(without alignments)
1331.818 Million cell updates/sec

Title: US-09-720-934-2
Perfect score: 5895
Sequence: 1 MAQFPFGSLDWAIVTVE.....QVGLFSPSYVKLTMDPSQ 1143

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3222919 seqs, 2451570024 residues
Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO spool_p/US09720934/runat_29072004_164339_1274/app_query.fasta_1.1287
-DB=Published Applications NA -QFMT=fastap -SUFFIX=jul29.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720934 @CGN_1_1_731 @runat_29072004_164339_1274
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

1	4615	78.3	3319	11	US-09-764-875-88	Sequence 88, Appl
2	4612	78.2	3466	16	US-10-158-057-33	Sequence 33, Appl
3	2837	48.1	5828	13	US-10-398-885A-15	Sequence 15, Appl
4	1728.5	29.3	2017	9	US-09-884-441-72	Sequence 72, Appl
5	1728.5	29.3	2017	10	US-09-907-969-72	Sequence 72, Appl
6	1728.5	29.3	2017	10	US-09-827-271-72	Sequence 72, Appl
7	1728.5	29.3	2017	15	US-10-198-053-72	Sequence 72, Appl
8	1089.5	18.5	2873	9	US-09-879-957-193	Sequence 193, Appl
9	1065	18.1	3746	11	US-09-764-868-125	Sequence 125, Appl
10	1045	17.7	4210	16	US-10-264-049-887	Sequence 176, Appl
11	757	12.8	2067	16	US-09-918-995-31258	Sequence 31258, A
12	752	12.4	503	10	US-09-764-881-55	Sequence 55, Appl
13	731.5	12.4	568	11	US-09-764-875-404	Sequence 404, Appl
14	731.5	12.4	568	13	US-09-764-881-55	Sequence 55, Appl
15	731.5	12.4	568	16	US-10-242-747-55	Sequence 55, Appl
16	731.5	12.4	568	16	US-10-158-057-127	Sequence 127, Appl
17	731.5	12.4	568	16	US-09-879-957-39	Sequence 39, Appl
18	634	10.8	747	9	US-10-342-887-1882	Sequence 1882, Ap
19	634	10.8	4053	13	US-10-172-118-1882	Sequence 1882, Ap
20	524	8.9	2442	9	US-09-964-899-24	Sequence 24, Appl
21	467	7.9	531	9	US-09-884-441-5	Sequence 5, Appl
22	467	7.9	531	10	US-09-907-969-5	Sequence 5, Appl
23	467	7.9	531	10	US-09-827-271-5	Sequence 5, Appl
24	467	7.9	531	15	US-10-198-053-5	Sequence 5, Appl
25	467	7.9	531	15	US-09-884-441-60	Sequence 60, Appl
26	456	7.7	480	10	US-09-907-969-60	Sequence 60, Appl
27	456	7.7	480	10	US-09-827-271-60	Sequence 60, Appl
28	456	7.7	480	15	US-10-198-053-60	Sequence 17127, A
29	456	7.7	480	15	US-09-864-761-17127	Sequence 26948, A
30	448	7.6	270	9	US-09-864-761-26948	Sequence 30453, A
31	448	7.6	286	9	US-09-864-761-30453	Sequence 36727, A
32	448	7.6	297	9	US-10-369-493-36727	Sequence 17146, A
33	444.5	7.5	3976	16	US-09-864-761-17146	Sequence 54, Appl
34	440	7.5	301	9	US-10-398-885A-1	Sequence 91, Appl
35	437	7.4	338	13	US-09-764-881-54	Sequence 54, Appl
36	436.5	7.4	955	11	US-09-764-875-91	Sequence 54, Appl
37	436.5	7.4	955	13	US-09-764-861-54	Sequence 54, Appl
38	436.5	7.4	955	16	US-10-242-747-54	Sequence 34, Appl
39	436.5	7.4	955	16	US-10-158-057-34	Sequence 17644, A
40	436.5	7.4	955	16	US-09-864-761-17644	Sequence 10314, A
41	427	7.2	263	9	US-09-864-761-10314	Sequence 50, Appl
42	427	7.2	480	9	US-09-764-881-50	Sequence 411, Appl
43	422.5	7.2	967	10	US-09-764-875-411	Sequence 50, Appl
44	422.5	7.2	967	11	US-09-764-875-411	
45	422.5	7.2	967	13	US-09-764-881-50	

ALIGNMENTS

RESULT 1
US-09-764-875-88
; Sequence 88, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ02
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 88
; LENGTH: 3319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-88

Alignment Scores:	0	Length:	3319
Pred. No.:	4615.00	Matches:	922
Score:	91.81%	Conservative:	9

Best Local Similarity: 90.93%
 Query Match: 78.29%
 DB: 11

Mismatches: 27
 Indels: 56
 Gaps: 4

US-09-720-934-2 (1-1143) x US-09-764-875-88 (1-3319)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
 DB 159 ATGGCTCAGTTTCCACACACTTTGGTGGCAGCTGGATATCTGGGCCATTAACGTAGAG 218
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProLysSerGlyPheIleThr 40
 DB 219 GAAAGAGCGAAGATGATCAGCAGTTCCATAGTTTAAAGCAATATCTGGATTTCATTACT 278
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 DB 279 GGTGATCAAGCTAGAACTTTTTCATCTGGTTTCACTCAACCTGTTTGTAGCACAG 338
 QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 DB 339 ATATGGCCTAGCTGACATGATTAATGATGGAAGATGATCAAGTGGAGTTTTCATA 398
 QY 81 AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 DB 399 GCTATGAATTTATCAAACTGAAAGCTACAGGATATCAGCTACCTCTGCACTTCGCCCT 458
 QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 DB 459 GTCATGAACAGACCAAGCTGTCTATTTCTAGCGCACCGACATTTGGTATGGAGGTATC 518
 QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 DB 519 GCCAGCATGCCACCGCTTACAGCTGTTCCTCCAGTCCCAATGGATCCATTCAGTTGTT 578
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 DB 579 GGAATGTCCTCAACCTAGTATCTCTGTTCCACAGCAGCTGTGCCCCCTGGCTAAC 638
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaIleThrLeuPro 180
 DB 639 GGGGCTCCCTCTGTATACACTCTGCTGCTGCAATTTGCTCATCTCGACGCCACATTGCCA 698
 QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 DB 699 AAGAGTCTCTCTTAGTAGATCTGCTCCAGGCTCACAACTAAACACTAAATTTACAAAG 758
 QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
 DB 759 GCACAGTCAATTTGATGTGCCAGTGTCCTCCAGTGGCAGAGTGGGCTGTTCTCTCAGTCA 818
 QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 DB 819 TCAAGACTGAATACAGCAATTAATCAATAGCTATGACAAACTATGATGGTGGACACTTA 878
 QY 241 Thr-----GlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGln 256
 DB 879 ACAGTTCCTGTTTAGTCCCAAGCAAGAACTATTCTTATGCACTCAAGTTTACACACAG 938
 QY 257 AlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAla 276
 DB 939 GCTCAGCTGGCTTCAATATGGAACTTTCTGCACTTGATCAAGATGGAAACTTACAGCA 998
 QY 277 GluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuPro 296
 DB 999 GAGGAATTTATCTGGCAATGACCTCATTTGATGTAGCTATGCTGTGGCAACCACTGGCA 1058
 QY 297 ProValLeuProProGluTyrIleProProSerPheArgValArgSerGlySerGly 316
 DB 1059 CCTGTCTCTCCCTCCAGAAATACATCTTCTTCTTAAAGAGATTCTGATCTGGCAGTGGT 1118
 QY 317 IleSerValIleSerSerThrSerValAspGlnArgLeuProGluGluProValLeuGlu 336
 DB 1119 ATATCTGTATGAAGCTCAACATCTGTAGATCAGAGGCTACAGAGAACCACTTTTAGAA 1178

QY 337 AspGluGlnGlnGlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGlu 356
 DB 1179 GATCAACAACAACAAATTTAGAAAAGAAATTAACCTGTACGTTTGAAGATAAGACGGGAG 1238
 QY 357 AsnPheGluArgGlyAsnLeuGluLysArgGlnAlaLeuLeuGluGlnGln 376
 DB 1239 AACTTTGAACTGGCAACCTGGAACTGGAGAAACGAAGGCAAGCTCTCTCTGGACACAG 1298
 QY 377 ArgLysGluGlnGlnArgLeuAlaGlnLeuGluArgAlaGluGlnGlnArgLysGluArg 396
 DB 1299 CGCAAGGAGCAGCAGCGCTGGCCAGCTGAGCGGGCGGCGAGGAGGAGGAGCGCT 1358
 QY 397 GluArgGlnGlnGlnArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGln 416
 DB 1359 GAGCGCCAGGAGCAAGCGCAAGACCAACTGGAACCTGGAAGCAACTGGAAGACGAG 1418
 QY 417 ArgGluLeuGluArgGlnArgGluGluArgArgLysGluIleGluArgArgGluAla 436
 DB 1419 CGGAGCTAGAACCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCT 1478
 QY 437 AlaLysArgGluLeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlu 456
 DB 1479 GCAAAAACGGGAACTTGAAGGCAACGACAACTTGTGAGTGGGAACGGAATCGAAGCAAGAA 1538
 QY 457 LeuLeuAsnGlnArgAsnLysGluGlnGluAspIleValLeuLysAlaLysLysLys 476
 DB 1539 CTACTAAATCAAGAAACAAAGAACAGAGGACATAGTTGTACTGAAAGCAAAAGAAAG 1598
 QY 477 ThrLeuGluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeu 496
 DB 1599 ACTTTGGAATTTGAATTAGAGCTCTAAATGATAAAGATCTCACTAGAGGGAACCTT 1658
 QY 497 GlnAspIleArgCysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSer 516
 DB 1659 CAAGATATCAGATGTGATTCACCCCAAGGCAAGAAATTTGAGAGCACAAACAATCT 1718
 QY 517 ArgGluLeuArgGluIleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnSerGlnGln 536
 DB 1719 AGAGAGTTGAGAAATTCGCAAAATCAACCATCTACAGCAACAATTTACAGGAATCTCAGCA 1778
 QY 537 MetLeuGlyValGluIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGln 556
 DB 1779 ATGCTTGGAGACTTATTCAGAAAAACAGATCTCAATGACCAATTAATAACAGTTTCA 1838
 QY 557 GlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGlu 576
 DB 1839 CAGAACAGTTTGCACAGAGATTCACTTGTACACTTAAAGAGCGCTTAGAAGCAAAAGAA 1898
 QY 577 LeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLys 596
 DB 1899 CTAGCTCGGACACCTTAGAGCACTGAGACCACTGGAGTGAAGTGGAGAAAGAACTAGATCAAAA 1958
 QY 597 LeuGlnGluIleAspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLys 616
 DB 1959 CTACAGGAGATTGATTTTCAATAATCAGCTGAAGGAACTAAGAGAAATAACAAATAG 2018
 QY 617 GlnGlnLeuGlnLysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGlu 636
 DB 2019 CAACAACCTCCAGAAACAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGCAAGAA 2078
 QY 637 ArgLysIleIleGluLeuGluLysGlnLysGluAlaGlnArgArgAlaGlnGluArg 656
 DB 2079 CGAAGATCATAGATTAGAAAAACAAAGAGAGCCCAAGAGCAGAGCTCAGGAAGG 2138
 QY 657 AspLysGlnTrpLeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeu 676
 DB 2139 GACAAGCAGTGGCTGGACATGTGCGCAGCAGGAGGAGGAGGATCAGAGACCAAGAAACTC 2198
 QY 677 HisGluGluGluLysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLys 696
 DB 2199 CACGAAGAGGAAAAAATCTGAAAAGGAGGAGAGTGTCTCAAAAAGAAAGGATGGCGAGGAAAA 2258

```
QY 697 GlyLysGlnGluAlaGlnAspLysLeuGlyArgLeuPheHisGlnGlnProAla 716
DB 2259 GCAAAACAGGAAGCAGACAGCTGGTGGCTTTCCATCAACACCAAGACAGCT 2318
QY 717 LysProAlaValGlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAla 736
DB 2319 AAGCCAGCTGTCCAGGACCTGGTCCACTGCAGAAAAGGTCACATTACCAATTTCTGCA 2378
QY 737 GlnGluAsnValLysValValTyrArgAlaLeuTyrProPheGluSerArgSerHis 756
DB 2379 CAGGAAATGTAAAGTGGTGTATTACCGGGACATGTACCCCTTTGAAATCCAGAGCCAT 2438
QY 757 AspGluIleThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlu 776
DB 2439 GATGAAATCATCTATCCAGCAGAGACATAGTATGATGATGAAAGCCAACTGGAGAA 2498
QY 777 ProGlyTrpLeuGlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAla 796
DB 2499 CCGCGTGGCTTGGAGGAGAAATTAAGGAAAGACAGGGTGGTTCCCTGCAACTATGCA 2558
QY 797 GluLysIleProGluAsnGlnValProAlaProValLysProValThrAspSerThrSer 816
DB 2559 GAGAAATCCAGAAATGAGGTTCCTCGCTCCAGTGAACCCAGTACTGATTCACATCT 2618
QY 817 AlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSer 836
DB 2619 GCCCTGCCCCCAACTGGCTTGGTGAGACCCCGCCCTTTGGCAGTAACTCTTCA 2678
QY 837 GluProSerThrThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThr 856
DB 2679 GAGCCCTCCAGCAGCCCTAATAACTGGGCGGACTTCAGCTCCACGTGGCCCAACAGCAG 2738
QY 857 AsnGluLysProGluThrAspAsnTrpAspAlaAlaGlnProSerLeuThrVal 876
DB 2739 AATGAGAAACAGAAACGGATACTGGGATGGATGGAGCCGCCCGCTCTCTCACCGTT 2798
QY 877 ProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySer 896
DB 2799 CCAAGTCCCGCCGAGTTAAGCAGAGGTCCGCTTTACTCCAGCCACGCGCTGGCTCC 2858
QY 897 SerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAlaLeu 916
DB 2859 TCCCGCTCTCTGGTGGTGGCAGCTGACTTCCTC-----CTC 2897
QY 917 TyrProTrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThr 936
DB 2898 CACCCC----- 2903
QY 937 ValLeuGluGlnGlnAspMetTrpTrpPheGlyGluValGlnGlnLysGlyTrpPhe 956
DB 2904 -----TCCATGAGACTAGGCCACATG 2924
QY 957 ProLysSerTyrValLysLeuIleSerGlyProIleArgLysSerThrSer----- 973
DB 2925 CAGCCCGGATGTGTCTTCTTCCCGACCCCTTCAGTGTTCACCTCCAGGCTTTTG 2984
QY 974 -----MetAspSerGlySerSerGluSerPro 982
DB 2985 CCCATGTGCGCCCTCGGCGTGGAGTGCCCTTCCTCCGAGCCCTAGTGTCAAGTCTCT 3044
QY 983 AlaSerLeuLysArgValAlaSerProAlaLysProVal 996
DB 3045 TCTCATCTCTCAAGGCCCATCTCAGATGCTGCCCCCTCTGTG 3086
```

RESULT 2

```
US-10-158-057-33
; Sequence 33, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ205C1
; CURRENT APPLICATION NUMBER: US/10/158,057
```

```
; CURRENT FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
```

```
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3194)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3465)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-158-057-33
```

```
Alignment Scores: 0 Length: 3466
Pred. No.: 4612.00 Matches: 921
Score: 91.72% Conservative: 9
Percent Similarity: 90.83% Mismatches: 28
Best Local Similarity: 78.24% Indels: 56
Query Match: 16 Gaps: 4
DB:
```

US-09-720-934-2 (1-1143) x US-10-158-057-33 (1-3466)

```
QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
DB 277 ATGGCTCAGTTTCCACACCTTTGGTGGCAGCTGGATATCTGGGCCATAAAGTGTAGAG 336
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
DB 337 GAAAGAGCGAAGCATGATCAGCAGTTTCCATAGTTTAAAGCCATATCTGGATTCAITACT 396
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
DB 397 GGTGATCAAGCTAGAACTTTTTCATCTGGGTACTCTCACTGTTTAGCACAG 456
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
DB 457 ATATGGCAGCTAGCTGACATGAATATGATGGAAGATGGATCAAGTGGAGTTTCCATA 516
QY 81 AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
DB 517 GCTATGAAACTTATCAAACTGAGCTACAAGGATATCAGCTACCTCTGCACTTCCCTCT 576
QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
DB 577 GTCATGAACAGCAACAGTTGCTATTCTTAGCCGACCAAGCATTTTGGTATGGAGGTATC 636
QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
DB 637 GCCAGATGCCACCGCTTACAGCTGTTGCTCCAGTGCATATGGGATCCATCCAGTTGTT 696
QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
DB 697 GGAATGTCTCAACCCCTAGTATCTTCTGTCTCCACAGCAGCTGTGCCCCCTGGCTAAC 756
QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
DB 757 GGGGCTCCCTGTTTATACAACTCTGCTGCATCTTGTCTCATCTCGACGACCATTTGCCA 816
QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
DB 817 AAGAGTCTTCTTTAGTAGTCTGGTCCAGGGTCCAACTAAACACACTAAATTTACAAAG 876
QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
DB 877 GCACAGTCAATTTGATGTGGCCAGTGTCCACCAAGGAGAGTGGGCTGTTCTCCTCAGTCA 936
QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
```

Db	937	TCAAGACTGAAATACAGGCAATATTCAATAGTCATGACAAACTATGATGGACACTTA	996	Db	2017	CTAGTCGGCAGCACCCTACGACACCACTGGATGAGTGGAGAAAGAACTAGATCAAAA	2076
QY	241	Thr-----GlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGln	256	QY	597	LeuGlnGluIleAspIlePheAsnAsnGlnLeuLeuArgGluIleHisAsnLys	616
Db	997	ACAGGTTCCTGTTTAGTCCCAAGCAGAACTATTCTTATCAGTCAAGTTTACACAG	1056	Db	2077	CTACAGGAGATTGATATTTCATTAATCAGCTGAAGAACTAAGAGAAATAACACAATAAG	2136
QY	257	AlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAla	276	QY	617	GlnGlnLeuGlnLysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGlnGlu	636
Db	1057	GCTCAGCTGGCTTCAATATGGAATCTTCTGACATTGATCAAGATGGAAACTTACAGCA	1116	Db	2137	CAACAACCTCCAGAACCAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAACAGAA	2196
QY	277	GluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuPro	296	QY	637	ArgLysIleIleLeuLeuGluLysGlnLysGluGluAlaGlnArgAlaGlnGluArg	656
Db	1117	GAGGAATTTATCCTGGCAATGCACCTCATTTGATGTAGCTATGTCTGGCCACCACTGCCA	1176	Db	2197	CGAAGATCATAGATTAGAAAAACAAAAGAGAGAGCCCAAGACGAGCTCCGAAAGG	2256
QY	297	ProValLeuProProGluTyrIleProProSerPheArgArgValArgSerGlySerGly	316	QY	657	AspLysGlnTyrLeuGluHisValGlnGlnGlnAspGluHisGlnArgProArgLysLeu	676
Db	1177	CCTGTCCCTGCCFCCAGAAATACATCCACCTCTTTTGAAGAGTTTCGATCTGGCAGTGT	1236	Db	2257	GACAGCAGTGGCTGGAGCATGTGCAGCAGGAGGACGAGCATCAGAGACCAAGAAACTC	2316
QY	317	IleSerValIleSerSerThrSerValAspGlnArgLeuProGluGluProValLeuGlu	336	QY	677	HisGlnGluGluLysLeuLysArgGluSerValLysLysLysAspGlyGluGluLys	696
Db	1237	ATATCTGTCAATAGCTCAACATCTGTAGATCAGAGGCTTACCAGAGAACCATGTTTAGAA	1296	Db	2317	CACGAGAGGAGAAACTGAAAGGAGGAGAGTGTCAAAAAGAGGATGGCGAGAAAAA	2376
QY	337	AspGluGlnGlnGlnLeuLysLysLeuProValThrPheGluAspLysLysArgGlu	356	QY	697	GlyLysGlnGluAlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAla	716
Db	1297	GATGAACAACAACAATTAGAAAGAAATTTACCTGTAACTGTTGAAGATAAGAGCGGAG	1356	Db	2377	GGCAACACAGGAAGCACAGACCAAGCTGGGTGGCTTTTCCATCAACACCAAGAACCGCT	2436
QY	357	AsnPheGluArgGlyAsnLeuGlnLeuLysArgArgGlnAlaLeuLeuGlnGln	376	QY	717	LysProAlaValGlnAlaProTyrSerThrAlaGluLysGlyProLeuThrIleSerAla	736
Db	1357	AACTTTGAACGTGGCAACTGGAACTGGAGAAACGAAAGCGCAAGCTCTCTCGAAACAGAG	1416	Db	2437	AAGCCAGCTGTCCAGGCACCTGTGTCCACTGCAGAAAAAGGTCCTACCTTACCATTTCTGCA	2496
QY	377	ArgLysGlnGlnGluArgLeuAlaGlnLeuGluArgAlaGluGlnArgLysGluArg	396	QY	737	GlnGluAsnValLysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHis	756
Db	1417	CGCAAGAGCAGGAGCGCTGGCCAGCTGGAGCGGGCGGAGCAGGAGGAAGAGCGT	1476	Db	2497	CAGGAATAATGAAAGTGGTGTATTACCGGGCAGCTGTACCCCTTTGAAATCCAGAGCCAT	2556
QY	397	GluArgGlnGlnGlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGln	416	QY	757	AspGluIleThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlu	776
Db	1477	GAGCGCCAGGAGCAGAGCGCAAAAGACAACTGGAACTGGAGAACTGGAAAGCAG	1536	Db	2557	GATGAATCACTATCCAGCCAGGAGACATAGTCATGTGTGGATGAAGCCAACTGGAGAA	2616
QY	417	ArgGluLeuGluArgGlnArgGluGluGluArgLysGluIleGluArgGluAla	436	QY	777	ProGlyTyrLeuGlyGlyGluLeuLysGlyThrGlyTyrPheProAlaAsnTyrAla	796
Db	1537	CGGAGCTTAGAACGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCT	1596	Db	2617	CCCGCTGGCTTGGAGGAGGATTAAGAGGAAAGACAGGCTGTCTCCCTGCACAACTATGCA	2676
QY	437	AlaLysArgGluLeuGluArgGlnArgGlnLeuGluTyrGluArgAsnArgArgGlnGlu	456	QY	797	GluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThrSer	816
Db	1597	GCAAAACGGGAACCTGAAAGCAGCAGCAACTTGAGTGGGACCGGAATCGAGGCAAGAA	1656	Db	2677	GAGAAATCCCGAATAATGAGTTCCCGCTCCAGTGAAACAGGACTGATTCACACTCT	2736
QY	457	LeuLeuAsnGlnArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLys	476	QY	817	AlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSer	836
Db	1657	CTACTAATCAAGAAACAAAGAACCAAGAGACATAGTTGTACTGAAAGCAAGAAAAAG	1716	Db	2737	GCCCTGCCCCCAACTGGCTTGGTGAGACCCCGCCCTTTGGCAGTAACTCTTCA	2796
QY	477	ThrLeuGluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeu	496	QY	837	GluProSerThrThrProAsnAsnTyrAlaAspPheSerSerThrTyrProThrSerThr	856
Db	1717	ACTTTGAATTTGAATTAGAGCTCTAATGATAAAGCATCACTAAGAGGAGAACTT	1776	Db	2797	GAGCCCTCCAGCAGCCCTTAATACTGGGCGGACTTCAGTCCAGTGGCCACCAGCAGCAG	2856
QY	497	GlnAspIleArgCysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSer	516	QY	857	AsnGluLysProGluThrAspAsnTyrAspAlaTyrAlaGlnProSerLeuThrVal	876
Db	1777	CAAGATATCAGATGCTGATTGACCACCAAGGCAAGAAATTTGAGAGCACAACAATCT	1836	Db	2857	AATGAGAAACAGAAACGGATTAAGTGGATGCTGGGAGCCCGAGCCCTCTCTCACCGCTT	2916
QY	517	ArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGlnGln	536	QY	877	ProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySer	896
Db	1837	AGAGAGTTGAGAAATGGCCGAATCACCATCTACAGCAACAATTTACAGGAATCTCAGCA	1896	Db	2917	CCAAGTCCCGCCAGTAAAGCAGAGTCCCGCTTTTACTCCAGCCACGGCCACTGGCTCC	2976
QY	537	MetLeuGlyArgLeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGln	556	QY	897	SerProSerProValLeuGlyGlnGlyValGluLysValGluGlyLeuGlnAlaGlnAlaLeu	916
Db	1897	ATGCTTGGAGACTTATCCAGAAAAACAGATCTCAATGACCAATTAAGCAAGTTTCAG	1956	Db	2977	TCCCGCTCTCTGTGTAGGCCAGCCGTACTCTCTC-----CTC	3015
QY	557	GlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGlu	576	QY	917	TyrProTyrArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThr	936
Db	1957	CAGAACAGTTTGACACAGAGATTCACTTTGTACACTTAAAGAGCCTTTAGAAGCAAAAGAA	2016	Db	3016	CACCCC-----	3021
QY	577	LeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLys	596	QY	937	ValLeuGlnGlnAspMetTyrTrpPheGlyGluValGlnGlyGlnLysGlyTyrPhe	956
				Db	3022	-----TCCATGAGACTAGGCCCATG	3042

QY 957 ProLysSerTyrValIysLeuIleSerGlyProIleArgLysSerThrSer----- 973
 Db 3043 CAGCCCGGATTGTCTCTCCCGGACCCCTGCAGTGTTCACACCTCGAGGCTTTTG 3102
 QY 974 -----MetAspSerGlySerGluSerPro 982
 Db 3103 CCCATGCTGGCCCTCGGCTGGAGTGCCTTCTCTCGAAGCCCTAGCTGTCAAGTCCT 3162
 QY 983 AlaSerLeuLysArgValAlaSerProAlaAlaLysProVal 996
 Db 3163 TCTCATCTTCAAGGSCCATCTCAGATGCTGNCCTCTCTGTG 3204

RESULT 3

US-10-398-885A-15
 ; Sequence 15, Application US/10398885A
 ; Publication No. US20040053282A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sugita, Yuji
 ; APPLICANT: Hashida, Ryoichi
 ; APPLICANT: Ogawa, Kaoru
 ; APPLICANT: Nagasu, Takeshi
 ; APPLICANT: Obayashi, Masaya
 ; APPLICANT: Saito, Hirohisa
 ; APPLICANT: Takahashi, Eiki
 ; TITLE OF INVENTION: Method of Testing For Allergic Diseases
 ; FILE REFERENCE: SHIMIZU-07907
 ; CURRENT APPLICATION NUMBER: US/10/398,885A
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: PCT/JP01/08937
 ; PRIOR FILING DATE: 2001-10-11
 ; PRIOR APPLICATION NUMBER: JP 2000-314093
 ; PRIOR FILING DATE: 2000-10-13
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 5828
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (7)..(5052)
 ; OTHER INFORMATION:
 ; US-10-398-885A-15

Alignment Scores:
 Pred. No.: 7,62e-207 Length: 5828
 Score: 2837.00 Matches: 627
 Percent Similarity: 64.01% Conservative: 175
 Best Local Similarity: 50.04% Mismatches: 252
 Query Match: 48.13% Indels: 200
 DB: 13 Gaps: 31

US-09-720-934-2 (1-1143) x US-10-398-885A-15 (1-5828)

QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
 Db 43 ATGGCTCAGTTTCCACAGCTATGATGAGGGCCCAACATGTGGCTATTACCTCTGAA 102
 QY 21 GluArgAlaLysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 103 GAACGTACTAAGCATGAGCAGGCGTTTGATAACCTCAACCTTCAGGAGGTTTACATAACA 162
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 163 GGTGATCAAGCAGCGTAATTTTCTCAATCAGCTGTCCGCGCCCTGTTTACTGAA 222
 QY 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 223 ATATGGGCTTTATCAGACCTAAACAGGATGGAAGATGGATCAGCAAGAGTTCTCCATA 282

QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 283 GCTATGAAACTCATCAAACTGAAGCTTCAAGGCCAACAGTTGCTGTGGTTCCTCCCTCCT 342
 QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
 Db 343 ATTATGAAGCAACCCCTATGTTTCTCCATTAATTTCTGCT---CGTTTTGAATGGGA 399
 QY 119 GlyIleAlaSerMetProLeuThrAlaValAlaProValProMetGly----- 135
 Db 400 -----AGCATGCCCATCTGTCCATTCCTCAGCCATTCCTCAGCTGCACCTATA 450
 QY 136 -----SerIleProValValGlyMetSerProThr 145
 Db 451 ACATCATTTGCTTCTCGCACTTCAGGGACCAACCTTCTCCCTTAATGATGCCCACTCCC 510
 QY 146 LeuValSerSerValProThrAlaAlaValProLeuAlaAsnGlyAlaProVal 165
 Db 511 CTAGTGCCTTCTGTGTAGCACATCATCATTAACCA-----AATGGAACCCGCACTCTC 561
 QY 166 IleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSerPhe 185
 Db 562 ATTCAGCCTTTACCC---ATTCTTATTTCTTCAACATTCCTCATGGGTCTCTTAT 618
 QY 186 Ser-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGlnSer 203
 Db 619 AGTCTGATGATGGAGGATTGGA-----GGTGCTAGTATACAGAAAGCGCAGTCT 669
 QY 204 ---PheAspValAlaSerValProPro----- 211
 Db 670 CTGATTGATTAGGATCTAGTAGCTCAACTCTCTCGACTCTTCACTCTCAGGAACTCA 729
 QY 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysTrpArgGln 727
 Db 730 CCCAAGACTGGGACCTCAGAGTGGGCGAGTCTCAGCCTACAGATTAAATATCGGCAA 789
 QY 228 LeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArgThr 247
 Db 790 AAATTTAATACTCTTGACAAAAGTATAGTGGATATCTCTCAGGTTTCAAGCTAGAAAT 849
 QY 248 IleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAsp 267
 Db 850 GCCCTCTCTCAGTCAAACTCTTCTCAAACTCAGCTGGCTACTATTGGGACTCTGGGTGAC 909
 QY 268 IleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIleAsp 287
 Db 910 GTTGATGGTGGGACAGCTAAAGCAGAGAGTTTATTCTTGCATGCGACCTTACTGAC 969
 QY 288 ValAlaMetSerGlyGlnProLeuProValLeuProGluTrpIleProSer 307
 Db 970 ATGGCCAAAGCTGGACAGCATTACCACTGACTTTTACCTCTCTGAGCTTGTCTCCATCT 1029
 QY 308 PheArgValAlaGlySerGlySerGlyIleSerValIleSerSerThrSerValAspGln 327
 Db 1030 TTC-----AGAGGGAGAAACAAATTGATTCATTAATGGAAC----- 1068
 QY 328 ArgLeuProGluGluProValIleLeuGluAspGluGlnGlnLeuGluLysLysLeuPro 347
 Db 1069 ---CTGCTTCTATATCAGAAATGCAAGAGAGGAGCCTCAG-----AAGAAATTACCA 1119
 QY 348 ValThrPheGluAspLysLeuAsnPheGluArgGlyAsnLeuGluLeuLys 367
 Db 1120 GTTACTTTTGAGGCAAAACGGAAAGCCCACTATGAGCGAGGGAACATGAGCTGGAAGA 1179
 QY 368 ArgArgGlnAlaLeuLeuGluGlnArgLysGluGlnGluArgLeuAlaGlnLeuGlu 387
 Db 1180 CGAGCCCAAGCTTGTATGAGCAGCACAACAAAGGGAGGAGCAGAACCTAAAGCCAGAA 1239
 QY 388 ArgAlaGluGlnGluArgLysGluArgGluArgGlnGluGlnGluArgLysArgGlnLeu 407
 Db 1240 AAGGAAGAGTGGGAACGAAACAGAGAGATTTACAGAACCAAGATGGAAGAAACACTT 1299
 QY 408 GluLeuGluLysGlnLeuGluLysGlnArgGluLeuGluArgGlnArgGluGluArg 427

1300	GAATAGAAAAACGCTTAGAGAACACCGGAATTTGGAGACACACGAGAGAGAGAGG	1359	2278	GATGAGATGAGCTTTAAATTTCTGGAGATATAATTCAGGTTGATGAAAAAACCGTAGAGAA	2337
428	ArgLysGluIleGluArgGluAlaAlaLysArgGluLeuGluArgGlnArgGlnLeu	447	777	ProGlyTyrLeuGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAla	796
1360	AGAAAGACATAGAAAGACGAGGACGACAAACAGAACTTGAACGACACGTCGTTA	1419	2338	CTGTGTTGCTTTATGTTAGTTTCAAGGAATTTTGGCTGTTTCCATGCAATATGTA	2397
448	GluTyrGluArgAsnArgGlnGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	467	797	GluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThrSer	816
1420	GAATGGAGAGATTCGGCGACAGAGAGCTTCTCAATCAAAAGAAATAGACAAAGAA	1479	2398	GAATAATGCCATCAAGTAA	2427
468	IleValValLeuLysAlaLysLysThrLeuGluPheGluLeuGluAlaLeuLeuLeu	487	817	AlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSer	836
1480	ATTGTCAAGCTTAACTTAAAGAAAGAAATCTTCATCTTGAGTTGGAACACTGAATGC	1539	2428	GCTGTATCTCCAAAGAGCCCTTACTTCTCTACAGTTTCTTATCTGCTACTCA	2484
488	LysLysHisGlnLeuGluGlyLysLeuGlnAspIleArgCysArgLeuThrGlnArg	507	837	GluProSerThrThrProAsnAsnTyrAlaAspPheSerSerThrTyrProThrSerThr	856
1540	AAACATCAGACGATCTCAGGACACTTCAGGATGTCGACTCAAAAGCAAACTCAAAAG	1599	2485	-----ACITCTCTGAACCCACTTCTCTCA	2508
508	GlnGluIleGluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHisLeu	527	857	AsnGluLysProGluThrAspAsnTyrAspAlaTyrAlaAlaGlnProSerLeuThrVal	876
1600	ACTGAGCTGAACTTCTGGATAAGCAGTGTGACTTGGAAATATATGAAATCAAGCAACT	1659	2509	AATCAACACGACATCAGTACTGATTATCAAAAT--GTATCTTTTCAACCTCAACTGTA	2565
528	GlnGlnGlnLeuGlnGlnSerGlnMetLeuGlyArgLeuIleProGluLysGlnIle	547	877	ProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySer	896
1660	CAACAGCACTTCAGGAATATCAGAAATAGCTTATCTATCTGTACTGAGACGAAATTA	1719	2566	AATACATCATGGCAG-----AAAAATCAGCCTTCTCTGAACTGTGCTCCCTGGA	2616
548	LeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuValThr	567	897	SerProSerProValLeuGlyGlnGlyGluLysValGluGlyLeuGlnAlaGlnAlaLeu	916
1720	TTAAATGAAGAATTAACATGCAGCTTCAGTAACACA--CCTGATTCAAGGGTCACT	1776	2617	TCTGTATCACTTATCATGACAGGACAGTGTAGAAACTTAAAGACACAGGCCCTT	2676
568	Leu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAspGln	585	917	TyrProTyrArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnValIleThr	936
1777	TTACTTCATATAAAATCATTAGAAAAGAGAAATATATGC--CAAGACTTAAAGAACAG	1833	2677	TGTTCTCGACTGCAAAAGAAAGATAACCACTTGAACCTTCTCAAAACATGATCATTTACT	2736
586	LeuAspGluValGlnLysGluThrArgSerLysLeuGlnGluIleAspIlePheAsnAsn	605	937	ValLeuGluGlnGlnAspMetTyrTyrPheGlyGluValGlnGlyGlnLysGlyTyrPhe	956
1834	TTAGATCTCTTGAAGAAGAACTGCATCTAAGCTGTCAAGATATGGAATCTTTTAAAT	1893	2737	GTCTTGAGCAGCAAGAAATTTGGTGTGTTGGGGAGTGTGATGGAGAGAGATGTTT	2796
606	GlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGlnLysGlnLysSerMet	625	957	ProLysSerTyrValLysLeuIleSerGlyPro-----	967
1894	CAACTAAGAACTGAGAGAAACCTACACACACAGCAGTTAGCCCTTGAACAGCTTTAT	1953	2797	CCCAATCTTATGTCAAGATCATCTCTGGGAGTGAAGTAAACCGGAAGAACACAGAGCT	2856
626	GluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIleGluLeu	642	968	-----IleArgLysSerThrSer-----	973
1954	AAGATCAACCTGACAAAGTTGAAGGAAATTCAGAGGAAAGATTAGAACTAATGCAGAA	2013	2857	TTGTATGACAGCTGTAAATAAGAAACCTACCTCGCAGCCTATTCACTGGAGAGAAATAT	2916
643	GluLysGlnLysGluGluAlaGlnArgAla--GlnGluArgAspLysGlnTyrLeu	661	973	-----	973
2014	AAGAACTTAGAAGATGAGCTGCAAGGAAAGCAAGCAAGCAAGAAAGAAACTTATGGAA	2073	2917	ATTGCACTTTATCCATATTCATTCAGTGTGGAACTCGAGATTTGACTTTTCACAGAGGTGAA	2976
662	GluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluLys	681	974	-----MetAspSerGlySerSerGlu-----	980
2074	GAATACTTAGAAGGAGAAAGAA-----GAAAAACAAAG	2109	2977	GAATATTTGGTGACCCAGAAAGATGGAGTGGTGGACAGCAAGTATTGGAGATAGAAAGT	3036
682	LeuLysArgGluGluSerValLysLysLysGlyGluGluLysGlyLysGlnGluAla	701	981	-----SerProAlaSerLeuLysArgValAlaSerProAlaAla	993
2110	CGACTCCAGGAGAAACCAAGAAATAATTCAGAAAGAGAGCAAGCAAGCTGAGAG	2169	3037	GGAATTTTCCATCAAACTATGTCAACCAAGGATCAAGAGAGTTT--GGAGTGTCTAG	3095
702	GlnAspLysLeuGlyArgPheHisGlnHisGlnGlnProAlaLysProAlaValGln	721	993	AlaProValValSerGlyGlu-----GluIleAlaGlnValIleAlaSerTyrThrAl	1011
2170	AAACAACGT-----AAGGATAAGGATACTTTGAAGCTGAGGAGAAAAACCT	2217	3096	CAAGTCTGGAGCATCAATAAATAAAACCTGAGATTGCTCAGGTAACTTCAGCATATGTTGC	3155
722	AlaProTyrSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnValLys	741	1011	aThrGlyProGlnGlnLeuThrLeuAlaProGlyGlnLeuIleLeuLeuArgLysLysAs	1031
2179	-----ValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHis	756	3156	TTCTGTTCTGAACAACTTAGCCTTGACCCAGGACAGTAAATTAATTAATTTCTAAAGAAAA	3215
2218	GAGACAGCTAGTGTGTTGTTGTAATATAGACATTTATACCTTTTGAACAGAGACCAT	2277	1031	nProGlyGlyTyrTyrGluGluLeuAlaArgGlyLysLysArgGlnIleGlyTyr	1051
757	AspGluIleThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlu	776	3216	TACAAAGTGGTGGTGGCAAGGAGATTACAGCCCAAGAGAAAGAAAGCAAGAGGATG	3275
			1051	pPheProAlaAsnTyrValLysLeuLeuSerProGlyThrSerLysIleThrProThrGln	1071
			3276	GTITCTCTGCACTCATGTTAATCTTTGGTCCCAAGTAGTGAAGAGACCACTGCCTT	3335

```
QY 1071 uProProLysSerThrAlaLeuAlaAlaValCysGlnValIleGlyMetTyrAspTyrTh 1091
Db 3336 TCATCCT-----GTAATGACAGTGAATGCTATGATGACTATGC 3374
QY 1091 rAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGlnIleAlaAsnValLeuAsnLy 1111
Db 3375 AGCAATAATAGATGAGTCACTAGTTCTCCAAAGGACAACTCATTAATGTTATGAACA 3434
QY 1111 sGluAspProAspTyrTrpLysGlyGluValAsnGlyGlnValGlyLeuPheProSerAs 1131
Db 3435 AGATGATCCTGATGTGGCGAAGAGAGATCAACGGGGTCACTGGTCTCTTTCCTTCAAA 3494
QY 1131 nTyrValLysLeuThrThrAspMetAspProSerGln 1143
Db 3495 CTACGTTAAGATGACGACGACTCAGATCCAAAGTCAA 3531

RESULT 4
US-09-884-441-72
; Sequence 72, Application US/0988441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF OVARIAN CANCER
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-72

Alignment Scores:
Pred. No.: 1,23e-122 Length: 2017
Score: 1728.50 Matches: 382
Percent Similarity: 70.63% Conservative: 99
Best Local Similarity: 56.09% Mismatches: 141
Query Match: 29.32% Indels: 59
DB: 9 Gaps: 18

US-09-720-934-2 (1-1143) x US-09-884-441-72 (1-2017)
QY 1 MetAlaGlnPheProThrPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 42 ATGGCTCAGTTTCCACAGCGATGAATGGAGGGCCAAATATGTTGGGCTATTACATCTGAA 101
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 102 GAACGTACTAAGCATGATGAACAGTTTGATTAACCTCAACCTTCAGAGGTTACATAACA 161
QY 41 GlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 162 GGTGATCAAGCCGCTACTTTTCTACAGTCAAGTCTGCGGCCCGCGTTTACGTGAA 221
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGCCCTTATCAGATCTCAACAGGATGGGAAGATGGACCAAGAGATTTCTCTATA 281
QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAAATCATCAAGTTTAAGTTGACGGCCCAACAGCTGCTGTAGTCTCTCTCTCT 341
QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAACAACACCCCTATGTTCTCTCCACTAATCTCTGCT---CGTTTGGGATGGGA 398
QY 119 GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal 132
Db 119 GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal 132
```

```
Db 399 -----AGCATGCCCAATCTGTCCATTCTATCAGCCATTGCTCCAGTTGCACCTATA 449
QY 133 -----ProMetGly-----SerIleProValValGlyMetSerPro 144
Db 450 GCAACACCCCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 509
QY 145 ThrLeuValSerSerValProThrAlaAlaValProLeuAlaAsnGlyAlaProPro 164
Db 510 CCCTAGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 560
QY 165 ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSer 184
Db 561 CTCATTTCAGCCCTTA---TCCATTCTCTTATCTCTCTCAACATGCTCTATGCTATCTCT 617
QY 185 PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGln 202
Db 618 TACAGCCTGATGATGGAGGATTTGCT-----GGTCTAGTATCCAGAGGCCAG 668
QY 203 Ser---PheAspValAlaSerValProPro----- 211
Db 669 TCTCTGATTGATTTAGGATCTAGTAGCTCAACTTCTCAACTGCTTCCCTCTCAGGGAAC 728
QY 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysTyrArg 226
Db 729 TCACCTAAGACAGGACCTCAGAGTGGCAGTTTCTCAGCCTTCAAGATTAAAGTATCGG 788
QY 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 789 CAAAAATTAATAGTCTAGACAAAGCATGACGCGATACCTCTCAGGTTTTCAGGTAGA 848
QY 247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 849 AATGCCCTTCTTTCAGTCAAACTCTCTCAAACTCAGTAGTACTATTTTGGACTCTGGCT 908
QY 267 AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle 286
Db 909 GACATCGATGGTGACGACAGTTGAAGCTGAAGAATTTATTCTGGCATGCACTCACT 968
QY 287 AspValAlaMetSerGlyGlnProLeuProValLeuProGluTyrIleProPro 306
Db 969 GACATGGCCAAAGCTGGACAGCCACTACCACTGACGTTGCTCCGAGCTTGTCCCTCCA 1028
QY 307 SerPheArgArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326
Db 1029 TCTTTC-----AGAGGGGAAAGCAAGTTGATTCTCTTAATGGAAC----- 1070
QY 327 GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnLeuGluLysLysLeu 346
Db 1071 -----CTGCTTTCATATCAGAAAAACAAGAAGAGAGCTCAG-----AGAAACTG 1118
QY 347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGlyAsnLeuGluGlu 366
Db 1119 CCAGTTACTTTTTCAGGACAAACGAAAGCCCACTATGAACGAGGAAACATGAGCTGGAG 1178
QY 367 LysArgArgGlnAlaLeuLeuGluGlnArgLysGluGlnGluArgLeuAlaGlnLeu 386
Db 1179 AAGCGCAGCCCAAGTGTGTGAGGACGACGACGAGGAGGCTGAACGAAAGCCGAGAA 1238
QY 387 GluArgAlaGluGlnGluArgLysGluArgGluGlnGluGlnArgLysArgGln 406
Db 1239 GAGAAGGAGAGTGGGAGCGGAAACAGAGAGAACTGCAAGAGCAAGATCGAAGAGCAG 1298
QY 407 LeuGluLeuLysGlnLeuGluLysGlnArgGluLeuGluArgGlnArgGluGlu 426
Db 1299 CTGGAGTTGGAGAAACGCTTTGGAGAAACAGAGAGCTGGAGAGACAGCGGAGGAAGAG 1358
QY 427 ArgArgLysGluIleGluArgArgGluAlaAlaLysArgGluLeuGluArgGlnArgGln 446
Db 1359 AGGAGAAAGGAGATAGAAACAGAGAGGCGACAAACAGGAGCTTGAGAGCAACACGCGT 1418
QY 447 LeuGluTrpGluArgAsnArgArgGlnGluLeuLeuAsnGlnArgAsnLysGluGlnGlu 466
Db 1419 TTGAATGGGAAAGACTCCGCTCGGAGGAGCTCTCTAGTCAGAGACCCAGGAGCAAGAA 1478
```

Qy 467 AspIleValValLeuLysAlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsn 486
Db 1479 GACATTGTCAGCTCAGAGCTCAGAAAGAAAGTCTCCACCTGGAGCTGGAAGCAGTGAAT 1538
Qy 487 AspLysLysHisGlnLeuGluGluGluLysLeuGlnAspIleArgCysArgLeuThrGln 506
Db 1539 GGAAACATCAGCAGATCTCAGGAGCAGACTACAAGATGTCCTCAATCAGAAACAAACAA 1598
Qy 507 ArgGlnGluLeuGluSerThrAsnLysSerArgGluLeuArgIleAlaGluLeuThrHis 526
Db 1599 AAGACTCAGCTAGAAAGTTTGGATAAACAGTGTGACCTGGAAATATGGAATCAACAA 1658
Qy 527 LeuGlnGlnGlnLeuGlnSerGlnGlnMetLeuGlyArgLeuLeuProGluLysGln 546
Db 1659 CTTCAACAGAGCTTAAGGATATCAAAATAGCTTATCTATCTGCTCCCTGAGAGCAG 1718
Qy 547 IleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuVal 566
Db 1719 CTAATTAACGAAAGAAATTAACAAATGAGCTCAGTCAACACA---CCTGATTCAGGGATC 1775
Qy 567 ThrLeu-----LysArgAlaLeuGluAlaLysGlnLeuAlaArgGlnHisLeuArgAsp 584
Db 1776 AGTTTACTTCATAAAAGTCTACGAAAGGAAAGAAATATATGC---CAAAGACTTAAGAA 1832
Qy 585 GlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGlnLeuIleAspIlePheAsn 604
Db 1833 CAATTAGATGCTCTTGAAGAAAGAACTGCATTAAGCTCTCAGAAATGATTCATTAAAC 1892
Qy 605 AsnGlnLeuLysGluLeuArgGluLeuIleHisAsnLysGlnGlnLeuLysGlnLysSer 624
Db 1893 AATCAGCTGAAGAACTCAGAGAAAGCTATAATACACAGCAGTGTAGCCCTTGAACAACTT 1952
Qy 625 MetGluAlaGluArgLeuLysGlnLysGlnGlnGluArgLysIleIleGluLeuGluLys 644
Db 1953 CATAAATCAACACGTGCAAAATTGAAGGAAATCGAAAGAAATGAGTATGAGCAAAAAA 2012
Qy 645 Gln 645
Db 2013 AAA 2015
RESULT 5
US-09-907-969-72
; Sequence 72, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-969-72
Alignment Scores: 1.23e-122 Length: 2017
Pred. No.: 1728.50 Matches: 382
Score:

Percent Similarity: 70.63% Conservative: 99
Best Local Similarity: 56.09% Mismatches: 141
Query Match: 29.32% Indels: 59
DB: 10 Gaps: 18
US-09-720-934-2 (1-1143) x US-09-907-969-72 (1-2017)
Qy 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 42 ATGGCTCAGTTTCCACAGCGATGAATGGAGGCCAAATATGTGGCTATTACATCTGAA 101
Qy 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 102 GAACGTACTAAGCATGATAAACAGCTTTGATAAACCTCAACCTTCAGGAGGTTTACATAACA 161
Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 162 GGTGATCAAGCCCTGACTTTTCTCAGTCAGTCTGCCGCCGCCGGTTTGTAGTGAA 221
Qy 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGCTTATCAGATCTGAACAGGATGGGAAGATGGACAGCAGAGAGTTCTCTATA 281
Qy 81 AlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAACCTCATCAAGTTAAAGTTGCAGGCGCAACAGCTGCCTGTAGTCTCCTCCT 341
Qy 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAACAAACCCCTATGTTCTCCACTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
Qy 119 GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal 132
Db 399 -----ACATGCCCAATCTCTCCATTCCATGCCATTCCTCCAGTTGGACCTATA 449
Qy 133 -----ProMetGly-----SerIleProValValGlyMetSerPro 144
Db 450 GCAACACCCCTGCT 509
Qy 145 ThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 510 CCCCTAGTGCCTCTCTGTTAGTACATCTCTCAATACCA-----AATGGAACTGCCAGT 560
Qy 165 ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSer 184
Db 561 CTCATTGAGCTTTA---TCCATTCTTATTCTTCTTCAACATTCCTCATGCTCATGCTCATCT 617
Qy 185 PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGln 202
Db 618 TACAGCTGATGATGGGAGGATTTGGT-----GGTGTAGTATATCCAGAGGCCAG 668
Qy 203 Ser---PheAspValAlaSerValProPro----- 211
Db 669 TCTCTGATTGATTGAGTCTAGTCTCACTTCCCTCACTGCTTCCCTCTCAGGGAAC 728
Qy 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysTyrArg 226
Db 729 TCACCTAAGACAGGACCTCAGAGTGGGAGCTTCTCAGCCTTCAAGATTAAAGTATCGG 788
Qy 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 789 CAAAAATTTAATAGTCTAGACAAAGGCATGAGCGGATACCTCTCAGGTTTTCAGACTAGA 848
Qy 247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 849 AATGCCCTTCTCTCAGTCAATCTCTCAAACTCAGTCTAGTACTATTGGACTTGGCT 908
Qy 267 AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle 286
Db 909 GACATCGATGGTGGAGCAGCTGAAAGCTGAAAGATTTATTCTGGCGATGCACCTCACT 968
Qy 287 AspValAlaMetSerGlyGlnProLeuProProValLeuProGluTyrIleProPro 306
Db 968

969 GACATGCCAAAGCTGGACAGCCACTACCACTGACGTTCCTCCGAGCTTGTCCTCCCA 1022

307 SerPheArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326

1029 TCITTC-----AGAGGGGAAGCAAGTTGATTCTGTTAATCGAACT----- 1070

327 GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnGlnLeuGluLysLysLeu 346

1071 -----CTGCCTTTCATACAGAAAACACACAGAAAGAGAGCCTCAG-----AAGAAACTG 1118

347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGluValAsnLeuGluGlu 366

1119 CCAGTTACTTTTCAGGACAAACCGCAAGCCAACTATGAACGAGAAACATATGGAGCTGGAG 1178

367 LysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGlnGluArgLeuAlaGlnLeu 386

1179 AAGCGACGCCAAGTTGATGGAGCAGCAGCAGGGAGGCGCTCAACGCCAAAGCCAGAAA 1238

387 GluArgAlaGlnGlnGluArgLysGluArgGluArgGlnGlnGluArgLysArgGln 406

1239 GAGAAAGAGAGTGGGAGCGGAAACACAGAGAACTGCAACAGCAAGAAATCGAAGAACAG 1298

407 LeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGluArgGlnArgGluGlu 426

1299 CTGGAGTTGGAGAAACCTCTGGAGAACACAGAGAGAGCTGGAGAGACAGCGGAGGAAGAG 1358

427 ArgArgLysGluIleGluArgArgGluAlaAlaLysArgGluLeuGluArgGlnArgGln 446

1359 AGAGAAAGAGAGATAGAAACGAGAGCGCAGCAAAACAGAGAGCTTGAGAGACAACGCCGT 1418

447 LeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGlnArgAsnLysGluGlnGlu 466

1419 TTAGAATGGCAAGACTCCGTCGCGCAGAGCTGCTCACTCAGAAGACAGCGGAACAAGAA 1478

467 AspileValValLeuLysAlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsn 486

1479 GACATTGTCCAGGTGAGCTCCAGAAAGAAAGTCTCCACTCGAACTGGAAGCAGTGAAT 1538

487 AspLysLysHisGlnLeuGluGluLysLeuGluAspileArgCysArgLysLeuThrGln 506

1539 GGAATAACATCAGCAGATCTCAGGCGAGCTACAGATGTCCTCAATCAGAAAGCAACACAA 1598

507 ArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArgileAlaGluIleThrHis 526

1599 AAGACTGAGCTAGAAAGTTTGGATAAACAGTGTGACCTCGAAATATTATGGAATCAAAACAA 1658

527 LeuGlnGlnGlnLeuGlnGluSerGlnGlnMetLeuGlyArgLeuIleProGluLysGln 546

1659 CTTTCAACAGAGCTTAAGGAATATCAAAAATAAGCTTATCTATCTGGTCCCTCGAAGACAG 1718

547 IleLeuAsnAspGlnLeuLysGlnValGlnGluAsnSerLeuHisArgAspSerLeuVal 566

1719 CTATTAACGAAAGAAATTAACATCGACTCAGTAACACA---CCTGATTACGGGATC 1775

567 ThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAsp 584

1776 AGTTTACTTCATAAAAAGTCATCAGAAAAAGGAAGAAATATATGCT---CAAGACTTAAAGAA 1832

585 GlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIleAspilePheAsn 604

1833 CAATTAGATGCTCTTGAAGAAAGAAACTGCATCTAAGCTCTCAGAAATATGGAATTCATTTAAC 1892

605 AsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGlnLysGlnLysSer 624

1893 AATCAGCTGAAGGAACCTCAGAAAGAGCTATTAATACACAGCAGATTAGCCCTTGAAACAATT 1952

625 MetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIleGluLeuGluLys 644

1953 CATTAATCAACCTGCACAAATTGAAGGAATTCAGAGAAATCGAAGAAAGAAATAGAGCAAAA 2012

645 Gln 645

2013 AAA 2015

Db 618 TACAGCTGATGATGGGAGGATTGGT-----GGTGCTAGTATCCAGAGGCCAG 668
Qy 203 Ser---PheAspValAlaSerValProPro----- 211
Db 669 TCTCTGATTGATTAGGATCTAGTAGCTCAACTTCCTCAACTGCTTCCCTCTCTCAGGGAAC 728
Qy 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysTrpArg 226
Db 729 TCACCTAAGACAGGAGCTCAGAGTGGGAGTCTCCTCAGCTTCCAGATTAAGTATCGG 788
Qy 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 789 CAAAAATTTAATAGTCTAGACAAAGGATGAGCGGATACCTCTCAGGTTTTCAGAGCTAGA 848
Qy 247 ThrLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 849 AATGCCCTTCTTCTAGTCAAACTCTCTCAAACTCAGCTAGTACTATTATGGACTCTGGCT 908
Qy 267 AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle 286
Db 909 GACATCGATGCTGACGACAGCTGAAGCTGAAGATTTATCTGGCGATGCACCTCACT 968
Qy 287 AspValAlaMetSerGlyGlnProLeuProProValLeuProProGluTrpIleProPro 306
Db 969 GACATGGCAAAGCTGACAGCCACTACCACTGAGCTTGCTCCCGAGCTGTGCCCTCCA 1028
Qy 307 SerPheArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326
Db 1029 TCTTTC-----AGAGGGGGAAGCAAGTTGATTTGTTAATGGAACCT----- 1070
Qy 327 GlnArgLeuProGluGluProValLeuGluAspGlnGlnGlnGlnLeuLysLysLeu 346
Db 1071 -----CTGCCTTCATATCAGAAACACAGAGAGAGCCCTCAG-----AAGAACTG 1118
Qy 347 ProValThrPheGluAspLysArgGluAsnPheGluArgGlyAsnLeuGluLeuGlu 366
Db 1119 CCAGTTACTTTTGGAGCAAAACCGAAACCACTATGAAACAGGAAACATGGAGCTGGAG 1178
Qy 367 LysArgGlnAlaLeuLeuGluGlnGlnArgLysGluGlnGluArgLeuAlaGlnLeu 386
Db 1179 AAGCGAGCCCAAGTGTGATGGAGCAGCAGCAGAGGGAGGCTGAACCGAAAGCCAGAAA 1238
Qy 387 GluArgAlaGlnGlnArgLysGluArgGluArgGlnGlnGlnGlnArgLysArgGln 406
Db 1239 GAGAAGGAGAGTGGAGCGGAGCAACAGAGAGACTGCAAGAGCAAGATGGAAGAGCAG 1298
Qy 407 LeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGluArgGlnArgGluGlu 426
Db 1299 CTGGAGTTGGAGAAACGCTTGGAGAAAACAGAGAGAGCTGGAGACACAGCGGGAGAGAG 1358
Qy 427 ArgArgLysGluIleGluArgGluAlaAlaLysArgGluLeuGluArgGlnArgGln 446
Db 1359 AGGAGAAGGAGTAGAAGAACGAGAGCGACCAACAAACAGGAGCTTGAGAGAACAGCCGT 1418
Qy 447 LeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGlnArgAsnLysGluGlnGlu 466
Db 1419 TTCAATGGGAAGAACTCCGTCGAGAGAGTGTCTCACTCAGAAAGACAGGCAACAGAA 1478
Qy 467 AspIleValValLeuLysAlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsn 486
Db 1479 GACATTTGACGCTGAGCTCCAGAAAGAAAGTCTCCACCTGGAACTCGAAGAGTGAAT 1538
Qy 487 AspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArgCysArgLeuThrGln 506
Db 1539 GGAACATCAGCAGATCTCAGGAGACTACAGATGTCCAAATCAGAAACAGAAACACAA 1598
Qy 507 ArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHis 526
Db 1599 AAGACTGAGCTAGAAGTTTGGATAAACAGTGTGACCTGGAAATTTATGGAATCAACAA 1658
Qy 527 LeuGlnGlnGlnLeuGlnSerGlnMetLeuGlyArgLeuIleProGluLysGln 546
Db 1659 CTTTCAACAGAGCTTAAGGAATATCAAAATTAAGCTTTATCTATCTGGTCCCTCGAAGCAG 1718

Qy 547 IleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuVal 566
Db 1719 CTATTAACCAAGAAATTAACATCAGCTCAGTACACACA---CCTGATTCAGGATC 1775
Qy 567 ThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAsp 584
Db 1776 AGTTTACTTCTATTAAGAGTCTATCAGAAAAGGAGAAATTTATGC---CAAAGACTTAAAGAA 1832
Qy 585 GlnLeuAspGluValGluLysGluThrArgSerLysLysGlnGlnGlnIleAspIlePheAsn 604
Db 1833 CAATTAGATCTCTTGAAGAAAGAACTGCATCTAAGCTCTCAGAAATGGATTTCAATTAAAC 1892
Qy 605 AsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuLysGlnLysSer 624
Db 1893 AATCAGCTGAAGAACTCAGAAAGCTATAATACACAGCAGTTAGCCCTTGAACAACCT 1952
Qy 625 MetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleLeuGluLeuGluLys 644
Db 1953 CATAAATCAACCTGACAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAA 2012
Qy 645 Gln 645
Db 2013 AAA 2015
RESULT 7
US-10-198-053-72
; Sequence 72, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10198.053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-053-72
Alignment Scores:
Pred. No.: 1,23e-122 Length: 2017
Score: 1728.50 Matches: 382
Percent Similarity: 70.63% Conservative: 99
Best Local Similarity: 56.09% Mismatches: 141
Query Match: 29.32% Indels: 59
DB: 15 Gaps: 18
US-09-720-934-2 (1-1143) x US-10-198-053-72 (1-2017)
Qy 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 42 ATGGCTCAGTTTCCACAGCCATGAATGGAGGGCAATATGTGGGCTATTACATCTGAA 101
Qy 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProfileSerGlyPheIleThr 40
Db 102 GAACGTACTAAGCATATAACAGTTTGATAACCTCAAACTTCAGGAGGTTTACATAACA 161
Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 162 GGTGATCAAGCCCTGACTTTTCTCAGTCAGTCTGCGGCCCGCCGTTTACCTGAA 221
Qy 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGGCTTATCAGATCTGAAACAGGATGGAGATGGAGCCAGCAGAGAGTTCTCTATA 281

QY 81 AlaMetLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAACCTCATCAAGTTAAAGTTGACGGGCAACAGCTGCTGTAGTCCCTCCCT 341
QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAACCAACCCCTATGTTCTCTCCACTATCTCTCT---CGTTTGGATGGGA 398
QY 119 GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal 132
Db 399 -----AGCATGCCCAATCTGTCCATTTCATCAGCCATTGCTCCAGTTGCACCTATA 449
QY 133 -----ProMetGly-----SerIleProValValGlyMetSerPro 144
Db 450 GCACACCCCTGTCTCTGCTACTTTCAGGACCACTATCTCCCTTAATGATGCCCTGCT 509
QY 145 ThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 510 CCCTAGTGCCTTCTGTAGTACATCTCATACCA-----AATGGAACTGCCAGT 560
QY 165 ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSer 184
Db 561 CTCATTACGCTTTA---TCCATTCTCTTATTCTTCTTCAACATTGCTCATGCTCATCT 617
QY 185 PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGln 202
Db 618 TACAGCTGATGATGGAGATTGGT-----GGTCTAGTATCCAGAGGCCAG 668
QY 203 Ser---PheAspValAlaSerValProPro----- 211
Db 669 TCTCTGATTGATTAGTATCTAGTCTCACTCACTTCTCACTTCTCCCTCTCAGGGAC 728
QY 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysTyrArg 226
Db 729 TCACCTTAAGACAGGGACCTCAGATGGGAGTTCTCTCAGCCTTCAAGATTAAAGTATCG 788
QY 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 789 CAAAAATTAAATAGTCTAGCAAAAGCATGAGCGGATACCTCTCAGGTTTTCAGGTAGA 848
QY 247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 849 AATGCCCTCTTTCAGTCAATCTCTCAAACTCAGTACTACTATTGGACTCTGGCT 908
QY 267 AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle 286
Db 909 GACATCGATGGTGCAGGACAGTTGAAAGCTGAGAAATTTATTCTGGCGATGCACCTCACT 968
QY 287 AspValAlaMetSerGlyGlnProLeuProValLeuLeuProGluTyrIleProPro 306
Db 969 GACATGGCCAAAGCTGGACAGCCTACCTGACCTGCTGCTCCCTGAGCTTGTCCCTCCA 1028
QY 307 SerPheArgArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValasp 326
Db 1029 TCTTTC-----AGAGGGGAAAGCAAGTTGATTCTGTTAATGGAAT----- 1070
QY 327 GlnArgLeuProGluProValLeuGluAspGluGlnGlnLeuGluLysLysLeu 346
Db 1071 -----CTGCCCTTCATATCAGAAAAACACAAAGAGAGAGAGCTCAG-----AAGAACTG 1118
QY 347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGluAsnLeuGluLeuGlu 366
Db 1119 CAGTTACTTTTGAGGACAAACGGAAGCACTATGAACGAGAAACATGAGCTGGAG 1178
QY 367 LysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGlnGlnArgLeuAlaGlnLeu 386
Db 1179 AAGCGACGCCAAGTGTGTGAGCAGCAGCAGCAGAGGAGGCTGAACGCAAGGCCAGAA 1238
QY 387 GlnArgAlaGluGlnGlnArgLysGluArgGluGlnGlnGlnArgLysArgGln 406
Db 1239 GAGAAGGAGAGAGTGGGAGCGGAACAGAGAGAACTGCAAGAGCAAGATGGAAGAGCAG 1298
QY 407 LeuGluLeuLysGlnGlnLysGlnArgGluLeuGluArgGlnGluGlu 426

Db 1299 CTGGAGTTGGAGAAACGCTTTGGAGAAACAGAGAGAGCTGGAGAGACAGCGGAGGAAG 1358
QY 427 ArgArgLysGluIleGluArgGluAlaAlaLysArgGluLeuGluArgGlnArgGln 446
Db 1359 AGAGAAAGAGATAGAAAGCAGAGAGGCGCAAAACAGGAGCTTGAGACACACGCGT 1418
QY 447 LeuGluTrpGluArgAsnArgArgGlnGluLeuLeuAsnGlnArgAsnLysGluGlnGlu 466
Db 1419 TTAGATGGGAAAGACTCCGTCGCGAGAGCTCTCAGTCAGAGACCGAGGACCAAGAA 1478
QY 467 AspIleValValLeuLysAlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsn 486
Db 1479 GACATTGTTCAGGCTGAGCTCCAGAAAGAAAGTCTCCACCTGGAACCTGGAAGCAGTGAAT 1538
QY 487 AspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArgCysArgLeuThrGln 506
Db 1539 GGAACATCAGCAGATCTCAGGACAGACTACAAGATGTCCAAATCAGAAAGCAACAA 1598
QY 507 ArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHis 526
Db 1599 AAGACTGAGCTAGAGCTTTTGGATAAACAGCTGTGACTGTGAAATTTATGGAATCAACAA 1658
QY 527 LeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArgLeuIleProGluLysGln 546
Db 1659 CTTCAACAGAGCTTAAGGAATATCAAAATAAGCTTATCTATCTGCTCCCTGAGAAGCAG 1718
QY 547 IleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuVal 566
Db 1719 CTATTAAACGAAGAATTAACATGCGAGCTCAGTAACACA---CCTGATTCAAGGATC 1775
QY 567 ThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAsp 584
Db 1776 AGTTTACTTCAATAAAAGTCATCAGAAAGAGGAATATTGTC---CAAGACTTAAGAA 1832
QY 585 GlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIleAspIlePheAsn 604
Db 1833 CAATTAGATGCTCTTGAAGAAAGAACTGCATCTTAAGCTCTCAGAAATGGATTCTTTAAC 1892
QY 605 AsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuLysGlnLysSer 624
Db 1893 AATCAGCTGAGGAACCTCAGAAAGCTATATATACACAGCAGTTAGCCCTTGAAACACTT 1952
QY 625 MetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIleGluLeuGluLys 644
Db 1953 CATAAATCAAAAGCTGACAAATTTGAAGAAATCGAAAGAAAGATTAGAGCAAAAAA 2012
QY 645 Gln 645
Db 2013 AAA 2015

RESULT 8
US-09-879-957-193
; Sequence 193, Application US/09879957
; Patent No. US20020034755A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; HOFFMAN, No. US20020034755A1h
; KAY, Brian K.
; FOWLES, Dana M.
; MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/879,957
 FILING DATE: 13-Jun-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/630,915
 FILING DATE: 03-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Mirock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 1101-174
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 193:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2873 bases
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 193:

US-09-879-957-193

Alignment Scores:

Pred. No.:	1.44e-73	Length:	2873
Score:	1089.50	Matches:	242
Percent Similarity:	55.88%	Conservative:	81
Best Local Similarity:	41.87%	Mismatches:	111
Query Match:	18.48%	Indels:	145
DB:	9	Gaps:	14

US-09-720-934-2 (1-1143) x US-09-879-957-193 (1-2873)

QY 636 GluArgLysIleIleGluLeu-----GluLysGlnLysGluGluAlaGlnArgArg 652
 DB 18 GAAGAGAAAGATTAGAACTAATGCAGAAAGAAAGAACTAGAAAGATGAGGCTGCAAGGAAA 77
 QY 653 Ala---GlnGluArgAspLysGlnTrpLeuGluHisValGlnGlnGluAspGluHisGln 671
 DB 78 GCAAGACGAAGAAAGAACTTATGGAAGAAATCTTGAAGAGGAGGAGAA-----131
 QY 672 ArgProArgLysLeuHisGluGluLysLeuLysArgGluGluSerValLysLysLys 691
 DB 132 -----GAAACAAAGCGACTCCAGGAGAAAGAAACACAGAAAGAAAT 176
 QY 692 AspGlyGluGluLysGlyLysGlnGluAlaGlnAspLysLeuGlyArgLeuPheHisGln 711
 DB 177 CAAGAGAGGAGGAAAGCGAAGAGCTGAGGAGAAACAACTGAGACAGCTAGTGTTC-----230
 QY 712 HisGlnGluProAlaLysProAlaValGlnAlaProTrpSerThrAlaGluLysGlyPro 731
 DB 230 -----230
 QY 732 LeuThrIleSerAlaGlnGluAsnValLysValValTyrTyrArgAlaLeuTyrProPhe 751
 DB 231 -----GTGAATTATAGAGCATTAACCCCTTT 257
 QY 752 GluSerArgSerHisAspGluIleThrIleGlnProGlyAspIleValMetValAspGlu 771
 DB 258 GAAGCAAGAAACCATGATGAGTGTAAATCTGGAGATATTAATTCAGGTGTATGAA 317
 QY 772 SerGlnThrGlyGluProGlyTrpLeuGlyGlyGluLeuLysGlyLysThrGlyTrpPhe 791
 DB 318 AAAACCGTAGGAGAACCTGGCTTTATGTTATGTTTCAAGGAAATTTGGCTGTTT 377
 QY 792 ProAlaAsnTyrAlaGluLysIleProGluAsnGluValProAlaProValLysProVal 811

DB 378 CCATGCATTTATGTAGAAAAAATGCCATCAAGTGA-----413
 QY 812 ThrAspSerThrSerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeu 831
 DB 414 -----AATGAAAAAGCTGTATCTCCAAAGAGCCCTTACTTCTCTCCTCAGTTCTTTA 467
 QY 832 AlaValThrSerSerGluProSerThrThrProAsnAsnTrpAlaAspPheSerSerThr 851
 DB 468 TCTGCTACCTCA-----ACTTCCTCT 488
 QY 852 TrpProThrSerThrAsnGluLysProGluThrAspAsnTrpAspAlaTrpAlaGln 871
 DB 489 GAACCACTTTCTTCAAAATCAACCAAGCATCAGTCAGTATCAAAAT---GTATCTTTT 545
 QY 872 ProSerLeuThrValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAla 891
 DB 546 TCAACCTTAACCTCTAAATACATCATGGCAG-----AAAAATCAGCCTTCACTCGAACT 599
 QY 892 ThrAlaThrGlySerSerProSerProValLeuGlyGlnGlyGluLysValGluGlyLeu 911
 DB 600 GTGTCCCTCGA---TCTGTATCACTTATTCATGGACAGGACAAAGTGGTAGAAACTTA 656
 QY 912 GlnAlaGlnAlaLeuTyrProTrpArgAlaLysLysAspAsnHisLeuAsnAsnLys 931
 DB 657 AAAGCACAGCCCTTTGTTCTCTGACTGCTGCAAGAAAGATACCACTTGAACCTTCTCAAAA 716
 QY 932 AsnAspValIleThrValLeuGluGlnGlnAspMetTrpTrpPheGlyGluValGlnGly 951
 DB 717 CATGACATTATTACTGTCTTGGAGCAGCAAGAAATTTGGTTTGGGAGGTGCATGGA 776
 QY 952 GlnLysGlyTrpPheProLysSerTyrValLysLeuLeuSerGlyPro-----967
 DB 777 GGAAGAGATGTTTCCCAAATCTTATGTCAAGATCATCTCTGGGAGTGGAAGTAAACGG 836
 QY 968 -----IleArgLysSerThrSer-----973
 DB 837 GAAGAACCAAGAGCTTTGTATGTCAGCTGTAAATGAAGAAACCTACTCGCAGCCTATTCA 896
 QY 973 -----973
 DB 897 GTTGGAGAGATATATTGCACTTTTATCCATATTCAAGTGTGGAACCTGGAGATTGACT 956
 QY 974 -----MetAspSerGlySerSerGlu--980
 DB 957 TTCACAGAAAGTGAAGAAATATTGTTGACCCAGAAAGATGAGAGTGTGGACAGGAAGT 1016
 QY 981 -----SerProAlaSerLeuLysArgVa 988
 DB 1017 ATTGGAGATGAAGTGGAAATTTTCCATCAAACTATGTCAACCAAGGATCAAGAGAGT 1076
 QY 988 lAlaSerProAlaAlaLysProValValSerGlyGlu-----GluIleAlaGlnVal 1006
 DB 1077 TTT-GGGAGTGTAGCAAGCTGGAGCATCAATAAATAAAACCTGAGATTGTCAGSTAAC 1135
 QY 1006 eAlaSerTyrThrAlaThrGlyProGluGluLeuThrLeuAlaProGlyGlnLeuIle 1026
 DB 1136 TTCAGCATATGTGCTTCTGTTCTGAAACACTTAGCCCTTGACCCAGGACAGATTATATT 1195
 QY 1026 uIleArgLysLysAsnProGlyGlyTyrTrpTrpGlyGluLeuGlnAlaArgGlyLys 1046
 DB 1196 AATTTCAAGAAAGAAATACAAAGTGGTGTGGCAAGAGAGATTACAGCCAGAGGAAAAA 1255
 QY 1046 sArgGlnIleGlyTrpPheProAlaAsnTyrValLysLeuSerProGlyThrSer 1066
 DB 1256 GCGACAGAAAGATGGTTTCTCTGCCAGTCATGTTAACTTTTGGGTCCAAGTAGTGAAG 1315
 QY 1066 sIleThrProThrGluProProLysSerThrAlaLeuAlaValCysGlnValIle 1086
 DB 1316 AGCCACACCTGCTTTCATCT-----GTATGTCAGGTGATTGC 1354
 QY 1086 yMetTyrAspTyrThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGlnIle 1106
 DB 1355 TATGTATGACTATGACGCAAAATATGAAGATGAGCTCAGTTTCTCCAGGAGCAACTCAT 1414

RESULT 10

US-09-764-868-125
; Sequence 125, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
; LENGTH: 4210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-125

Alignment Scores:
Pred. No.: 5,89e-70 Length: 4210
Score: 1045.00 Matches: 230
Percent Similarity: 55.51% Conservative: 72
Best Local Similarity: 42.28% Mismatches: 104
Query Match: 17.73% Indels: 140
DB: 9 Gaps: 11

US-09-720-934-2 (1-1143) x US-09-764-868-125 (1-4210)

Qy	666	GlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGluLysLeuLysArgGlu	685
Db	492	AGAAAGGAGGAAGAAACAAACAAAGCGCTCCAGGAGAGAAACAAACACAGAAAAA	551
Qy	686	GluSerValLysLysAspGluGluLysGlyLysGlnGluAlaGlnAspLysLeu	705
Db	552	TT-CAAGAAAGAGGAACGAAAGCTGAGGAGAAACAAAGTGGAGACAGCTAGT	601
Qy	706	GlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaValGlnAlaProT	725
Db	601	-----	601
Qy	726	ThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnValLysValTyr	745
Db	602	-----GTTTGTGTAATTAT	616
Qy	746	ArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGlnProGly	765
Db	617	AGAGCATTTATACCCCTTTGAAGCAAGAACCATGATGATGAGTGGTAAATTC	766
Qy	766	IleValMetValAspGluSerGlnThrGlyGluProGlyTyrLeuGlyGlyGluLys	785
Db	677	ATAATTCAGTTGTATGAAGAAACCCGTAGAGAACCTGGTTGGCTTTATGTTT	736
Qy	786	GlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIleProGluAsnGluVal	805
Db	737	GGAAATTTGGCTGTTTCCATGCAATATGATAGAAAAATGCCATCAAGTAA	790
Qy	806	AlaProValLysProValThrAspSerThrSerAlaProAlaProLysLeuAla	825
Db	791	-----AATCAAAAGCTGTATCTCCAAAGAGGCCCTTACT	826
Qy	826	GluThrProAlaProLeuAlaValThrSerSerGluProSerThrThrProAsn	845
Db	827	CCTCTACAGTTCTTTATCTGCTACCTCA	856
Qy	846	AlaAspPheSerSerThrTyrProThrSerThrAsnGluLysProGluThrAsp	865
Db	857	-----ACTTCTCTGAAACACTTTCTTCAATCAACACGACGATCATGCTGAT	907
Qy	866	AspAlaTyrAlaAlaGlnProSerLeuThrValProSerAlaGlyGlnLeuArg	885
Db	908	CAAAAT---GTATCTTTTCAACCTAACTGTAAATACATCATGTCAG-----	958

Qy	886	SerAlaPheThrProAlaThrAlaThrGlySerSerProSerProValLeuGlyGln	905
Db	959	TCAGCCTTCACTCGAACTGTGTCCCTGGA---TCTGTATCACCTATTTCATGCA	1015
Qy	906	GluLysValGluGluLeuGlnAlaGlnAlaLeuTyrProTyrAlaLysLysAsp	925
Db	1016	CAAGTGTAGAAACTTAAAGACACAGGCCCTTTGTTCTCTGAGCTGCAAGAA	1075
Qy	926	HisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMet	945
Db	1076	CACITGAACTTCTCAAAACATGACATTTATCTCTCTGGAGCAGCAAGAAAT	1135
Qy	946	PheGlyGluValGlnGlnLysGlyTyrPheProLysSerTyrValLysLeuIle	965
Db	1136	TTTGGGAGGTGCATGAGGAGAGGATGGTTTCCAAATCTTATGTCAGATCAT	1195
Qy	966	GlyPro-----IleArgLysSer	971
Db	1196	GGGAGTGAAGTAAACCGGAAGAACCAAGAGCTTTGTATGAGCTGTAAATA	1255
Qy	972	ThrSer-----	973
Db	1256	ACCTCGCAGCCTATTTCAGTTGGAGAGAAATATATTGCACCTTTATCCAT	1315
Qy	974	-----MetAs	975
Db	1316	GAACCTGGAGATTTGACTTTTCCACAGAGGTGAAGAAATATTGTGTGCCCA	1375
Qy	975	pSerGlySerSerGlu-----SerPr	982
Db	1376	GAGTGTGGCAGAGAGATTTTGGAGATAGAAAGTGAATTTTCCATCAAACT	1435
Qy	982	oAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGly	1000
Db	1436	CAAAGGATCAAGAGATTTT-GGGAGTGTAGCAGCTCTGGAGCATCAATAA	1494
Qy	1001	-GluIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnLeu	1020
Db	1495	TGATATTGCTCAGGTAACCTCAGCATATGTTGCTTCTGGTTCTGAACAACT	1554
Qy	1020	aProGlyGlnLeuIleLeuIleArgLysLysAsnProGlyGlyTyrProGlu	1040
Db	1555	ACCAAGCAGATTAATTAATTTCTAAAGAAAAATCAAGTGGTGGTGGCAGG	1614
Qy	1040	uGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrVal	1060
Db	1615	ACAGCCAGAGAAAAAGCCAGAGAGGATGGTTTCTCCAGTCATGTTAACT	1674
Qy	1060	uSerProGlyThrSerLysIleThrProThrGluProProLysSerThrAla	1080
Db	1675	GGGTCCAAGTAGTGAAGAGCCACACCTGCTTTCATCT	1714
Qy	1080	aValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAsp	1100
Db	1715	-GTATGTCAAGTGATGCTATGATGACTATGCGAGAAATTAATGAATGAG	1773
Qy	1100	eAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrP	1120
Db	1774	CTCCAAGGACAACTCATTAATGATTAAGCAAAAGATGATCTGATGTTG	1833
Qy	1120	uValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThr	1140
Db	1834	GATCAACGGGTGACTGGTCTCTTCTTCAAACTACGTTAAGATGACGAC	1893
Qy	1140	pProSerGln	1143
Db	1894	TCCAAGTCAA	1903

RESULT 11
US-10-264-049-887
; Sequence 887, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:

; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 887
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2058)..(2058)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2063)..(2063)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-887

Alignment Scores:
Pred. No.: 2,57e-48 Length: 2067
Score: 757.00 Matches: 141
Percent Similarity: 97.95% Conservative: 2
Best Local Similarity: 96.58% Mismatches: 3
Query Match: 12.84% Indels: 0
DB: 16 Gaps: 0

US-09-720-934-2 (1-1143) x US-10-264-049-887 (1-2067)

QY 998 SerGlyGluGluLeuAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnLeu 1017
DB 53 TCTCAAAATAAATTTGCCCCAGGTTATTGCTCATACACGCCACCGCCCGCCGAGCCTC 112
QY 1018 ThrLeuAlaProGlyGlnLeuIleLeuIleArgLysLysAsnProGlyGlyTrrPrrGlu 1037
DB 113 ACTCTCGCCCTGGTCAGCTGATTTTGATCCGAAAGAACCCAGGTGGATGGTGGAA 172
QY 1038 GlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTrrPheProAlaAsnTyrVal 1057
DB 173 GGAGAGCTCAAGCAGCTGGGAAAGAGCCAGATAGGCTGGTTCAGCTAATTATGTA 232
QY 1058 LysLeuLeuSerProGlyThrSerLysIleThrProThrGluProProLysSerThrAla 1077
DB 233 AAGCTTTCTAAGCCCTGGGACGAGCAAAATCACTCCAAACAGAGCCACCTAAGTCAACAGCA 292
QY 1078 LeuAlaAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGlu 1097
DB 293 TTAGCGGAGTGTGCCAGGTGATTTGGATGTAGACTACACCGCGCAGAAATGACCATGAG 352
QY 1098 LeuAlaPheAsnLysGlyGlnIleLeuAsnValLeuAsnLysGluAspProAspTrrPrr 1117
DB 353 CTGGCTTCAACAGGCGCCAGATCATACAGTCTCTCAACAGGAGACCTGACTGGTG 412
QY 1118 LysGlyGluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThr 1137
DB 413 AAAGGAGAGTCAATGGACAGTGGGGCTCTTCCCATCAATATGTGAAGCTGACCACA 472
QY 1138 AspMetAspProSerGln 1143
DB 473 GACATGGAGCCAGCCAG 490

RESULT 12

US-09-918-995-31258

; Sequence 31258, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31258
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(503)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31258

Alignment Scores:
Pred. No.: 1.05e-48 Length: 503
Score: 752.00 Matches: 152
Percent Similarity: 98.09% Conservative: 2
Best Local Similarity: 96.82% Mismatches: 3
Query Match: 12.76% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2 (1-1143) x US-09-918-995-31258 (1-503)

QY 276 AlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeu 295
DB 33 TCAGAGAAATTTATCTGGCAATGCACCTCATTTGATGATGCTGTGCCCAACCACTG 92
QY 296 ProProValLeuProGluTyrIleProProSerPheArgArgValArgSerGlySer 315
DB 93 CCACCTGTCTGCTCCAGAAATACATTCACCTCTCTTTAGAGAGTTGATCTGGCAGT 152
QY 316 GlyIleSerValIleSerThrSerValAspGlnArgLeuProGluGluProValLeu 335
DB 153 GGTATATCTGTATAGCTCAACATCTGTAGATCAGAGGCTACAGAGAACCACTTTTA 212
QY 336 GluAspGluGlnGlnGlnLeuGluLysLysLeuProValThrPheGluAspLysLysArg 355
DB 213 GAAGATGAACACACAAATTAGAAAAGAAATACCTGTAAACGTTTGAAGATAAGAGCGG 272
QY 356 GluAsnPheGluArgGlyAsnLeuGluLeuLysArgGlnAlaLeuLeuGluGln 375
DB 273 GAGAACTTTGAACGTGGCAACCTGGAACTGGAGAACGAGGCAAGCTCTCTGGAACAG 332
QY 376 GlnArgLysGluGlnGluArgLeuAlaGlnLeuGluArgAlaGluGlnArgLysGlu 395
DB 333 CAGCCCAAGAGCAGAGCGCTGGCCCACTGGAGCGCGGAGCAGGAGGAGGAG 392
QY 396 ArgGluArgGlnGlnGlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLys 415
DB 393 CGTGAGCGCCAGAGCAGAGCGCATAGACAACTGGAACTGGAGATGCAACTGGGAAAG 452
QY 416 GlnArgGluLeuGluArgGlnArgGluGluGluArgArgLysGluLeuGlu 432
DB 453 CAGCGGAGCTAGAACGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 503

RESULT 13

US-09-764-881-55

; Sequence 55, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207

Publication No. US20040018969A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PJ202
 CURRENT APPLICATION NUMBER: US/09/764,875
 CURRENT FILING DATE: 2001-01-17
 Prior application data removed - consult PALM or file wrapper
 NUMBER OF SEQ ID NOS: 1249
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 404
 LENGTH: 568
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (481)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (536)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (556)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (562)
 OTHER INFORMATION: n equals a,t,g, or c
 US-09-764-875-404

CURRENT APPLICATION NUMBER: US/09/764,881
 CURRENT FILING DATE: 2001-01-17
 Prior application data removed - refer to PALM or file wrapper
 NUMBER OF SEQ ID NOS: 192
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 55
 LENGTH: 568
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (481)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (536)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (556)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (562)
 OTHER INFORMATION: n equals a,t,g, or c
 US-09-764-881-55

Alignment Scores:
 Pred. No.: 4,53e-47 Length: 568
 Score: 731.50 Matches: 151
 Percent Similarity: 92.07% Conservatives: 0
 Best Local Similarity: 92.07% Mismatches: 12
 Query Match: 12.41% Indels: 2
 DB: 10 Gaps: 0

US-09-720-934-2 (1-1143) x US-09-764-881-55 (1-568)
 QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
 Db 78 ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCTGGATATCTGGGCCATAACTGTAGAG 137
 QY 21 GluArgAlaLysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 138 GAAAGAGCGA--CATGATCAGCAGTTCATAGTTTAAAGCCAATATCTGGATTCAATTACT 195
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 196 GGTGATCAAGCTAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
 QY 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 256 ATATGGGCACTAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 316 GCTATGAAACTTATCAAACTGAAGCTACAAGGATATCAGTACCTCTGCATCTCCCTCT 375
 QY 101 ValMetLysGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 Db 376 GTTCATGAAACACCAACCACTTCTATTTCTAGGCAACCACTTCTAGGCAACCACTTCTAGG 435
 QY 121 AlaSerMetProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 436 GCCAGCAACCAACCACTTACAGCTGTGTCTCCAGTGCCTGCAATGGGNCCTATCCAGTTGT 495
 QY 141 -GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAs 160
 Db 496 GGGAAATGTCTCCAAACCTAGTATCTTCTGTCTCCACAGCANTGTGCCCCCTGGGTAAA 555
 QY 160 nglyAlaPro 163
 Db 556 NGGGGTNCCT 565

Alignment Scores:
 Pred. No.: 4,53e-47 Length: 568
 Score: 731.50 Matches: 151
 Percent Similarity: 92.07% Conservatives: 0
 Best Local Similarity: 92.07% Mismatches: 12
 Query Match: 12.41% Indels: 2
 DB: 11 Gaps: 0

US-09-720-934-2 (1-1143) x US-09-764-875-404 (1-568)
 QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
 Db 78 ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCTGGATATCTGGGCCATAACTGTAGAG 137
 QY 21 GluArgAlaLysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 138 GAAAGAGCGA--CATGATCAGCAGTTCATAGTTTAAAGCCAATATCTGGATTCAATTACT 195
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 196 GGTGATCAAGCTAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
 QY 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 256 ATATGGGCACTAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 316 GCTATGAAACTTATCAAACTGAAGCTACAAGGATATCAGTACCTCTGCATCTCCCTCT 375
 QY 101 ValMetLysGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 Db 376 GTTCATGAAACACCAACCACTTCTATTTCTAGGCAACCACTTCTAGGCAACCACTTCTAGG 435
 QY 121 AlaSerMetProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 436 GCCAGCAACCAACCACTTACAGCTGTGTCTCCAGTGCCTGCAATGGGNCCTATCCAGTTGT 495
 QY 141 -GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAs 160
 Db 496 GGGAAATGTCTCCAAACCTAGTATCTTCTGTCTCCACAGCANTGTGCCCCCTGGGTAAA 555
 QY 160 nglyAlaPro 163
 Db 556 NGGGGTNCCT 565

Db 556 NGGGGTNCCC 565

RESULT 15

US-09-764-881-55

; Sequence 55, Application US/09764881

; Publication No. US20020086821A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ07

; CURRENT APPLICATION NUMBER: US/09/764,881

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 192

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 55

; LENGTH: 568

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (481)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (536)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (556)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (562)

; OTHER INFORMATION: n equals a,t,g, or c

US-09-764-881-55

Alignment Scores:

Pred. No.:	4.53e-47	Length:	568
Score:	731.50	Matches:	151
Percent Similarity:	92.07%	Conservative:	0
Best Local Similarity:	92.07%	Mismatches:	12
Query Match:	12.41%	Indels:	2
DB:	13	Gaps:	0

US-09-720-934-2 (1-1143) x US-09-764-881-55 (1-568)

QY	1	MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu	20
Db	78	ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCCTGGATATCTGGGCCATAACTGTAGAG	137
QY	21	GluAtgAlaIysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr	40
Db	138	GAAGAGCGCA--CATGATCAGCAGTTCCTAGTGTAAAGCCATATCTGGATTCTACT	195
QY	41	GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln	60
Db	196	GGTGATCAAGCTAGAAACCTTTTTCATCTGGGTACCTCAACCTGTTTTCAGCAG	255
QY	61	IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle	80
Db	256	ATATGGGCACCTAGCTGACATGAATAATGATGAAGAATGGATCAAGTGGAGTTTCCATA	315
QY	81	AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuPro	100
Db	316	GCTATGAAACTTATCAAACTGAAGCTACAGGATATCAGTACCCCTCTGCACCTCCCCCT	375
QY	101	ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle	120
Db	376	GTATGAACACGACCAACAGTTCCTATTTCTAGGCCACAGCATTTGGTATGGGAGGTATC	435
QY	121	AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal	140
Db	436	GCCAGCAAGCCACCGCTTACAGCTGTGCTCCAGTGCCAATGGGNGCCATTCAGTTGTT	495
QY	141	-GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProLeuAlaAs	160

Db 496 GGGATGTCTCCAACCCCTAGTATCTTCTGTCCACAGCANTGTGCCCCCTGGCTAAA 555

QY 160 nGlyAlaPro 163

Db 556 NGGGGTNCCC 565

Search completed: August 3, 2004, 10:05:36

Job time : 4280 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 2, 2004, 17:25:24 ; Search time 909 Seconds
(without alignments)

5341.794 Million cell updates/sec

Title: US-09-720-934-2

Perfect score: 5895

Sequence: 1 MAQFTPPGGSLDIWAIVE.....QVGLPFSNVKLTMDPSQ 1143

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cpn2.1/uspto.spool.p/US09720934/runat.29072004.164337.1209/app.query.fasta.1.1287
-DB=N_Geneseq.25Jan04 -QFMT=fastap -SUFFIX=jul29.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -NAXLEN=2000000000
-USER=US09720934 @CGN 1.1.708 @runat.29072004.164337.1209 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5895	100.0	5199	2	AZ34570 Human SH3
2	5849.5	99.2	5195	2	AZ34572 Human SH3
3	5837	99.0	5458	2	AZ34571 Human SH3
4	5441	92.3	7435	5	AAS84763 DNA encod
5	5437.5	92.2	3723	3	AZ39009 Mouse Ese
6	5437.5	92.2	5144	3	AZ39025 Mouse Ese
7	5437.5	92.2	5738	3	AZ39024 Mouse Ese
8	5408.5	91.7	5082	3	AZ39008 Mouse Ese

9	4615	78.3	3319	4	ABK43498	Abk43498 DNA encod
10	4612	78.2	3466	4	AAI63825	Aai63825 Human pol
11	3344	56.7	3231	2	AZ34574	Aaz34574 Human SH3
12	3246	55.1	2079	2	AZ34573	Aaz34573 Human SH3
13	3209	54.4	2131	4	AAK94139	Aak94139 Human ful
14	3133.5	53.2	2874	5	AAS84762	Aas84762 DNA encod
15	2962.5	50.3	2131	4	AAH16578	Aah16578 Human CDN
16	2837	48.1	5828	6	AAI47247	Aai47247 Allergic
17	2813.5	47.7	6103	4	AAK52332	Aak52332 Human pol
18	2779	47.1	4625	3	AZ39010	Aaz39010 Mouse Ese
19	2779	47.1	4975	3	AZ39027	Aaz39027 Mouse Ese
20	2779	47.1	6014	3	AZ39026	Aaz39026 Mouse Ese
21	2758	46.8	3593	3	AZ39011	Aaz39011 Mouse Ese
22	2740.5	46.5	4447	4	AAS02055	Aas02055 DNA encod
23	2436	41.3	1389	2	AAT39795	Aat39795 Human SH3
24	2388	40.5	1676	4	AAK94611	Aak94611 Human CDN
25	1728.5	29.3	2017	3	AAJ69762	Aaj69762 Human ova
26	1728.5	29.3	2017	3	ABN72656	Abn72656 Ovarian c
27	1728.5	29.3	2017	8	ADA08821	Ada08821 Human ova
28	1671.5	28.4	3981	4	ABL01995	Abi01995 Drosophil
29	1468	24.9	7225	4	AAH01994	Aah01994 Human CDN
30	1231	20.9	1329	4	AAH15260	Aah15260 Human CDN
31	1110	18.8	677	7	ACA57641	ACA57641 Human adi
32	1093.5	18.5	2873	2	AAT39799	Aat39799 Human cto
33	1065	18.1	3746	4	ABK43586	Abk43586 DNA encod
34	1045	17.7	4210	4	AAS27090	Aas27090 cDNA enco
35	1045	17.7	4210	9	ADB93268	Adb93268 Human CDN
36	994	16.9	877	4	AAK93179	Aak93179 Human CDN
37	994	16.9	877	4	AAK91610	Aak91610 Human CDN
38	846	14.4	831	4	AAH03435	Aah03435 Human pol
39	796	13.5	548	4	AAI80000	Aai80000 Human pol
40	757	12.8	2067	6	ABQ55007	Abq55007 Human ova
41	752	12.8	503	8	ACH44046	Ach44046 Human foe
42	731.5	12.4	568	4	AAI63919	Aai63919 Human pol
43	731.5	12.4	568	4	AAK31621	Aak31621 cDNA enco
44	731.5	12.4	568	4	ABK43814	Abk43814 DNA encod
45	634	10.8	747	2	AAT39796	Aat39796 Human SH3

ALIGNMENTS

RESULT 1
AAZ34570
ID AAZ34570 standard; cDNA; 5199 BP.

XX AAZ34570;

DT 01-FEB-2000 (first entry)

XX Human SH3DIA cDNA.

XX SH3DIA gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FT CDS 208..3642
FT /*tag= a

XX WO9953062-A2.

XX 21-OCT-1999.

XX 16-APR-1999; 99WO-US008371.

XX 16-APR-1998; 98US-0082007P.

XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.

PI Korenberg JR, Chen X;

XX WPI; 1999-633829/54.

DR P-PSDB; AAY32154.

XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.

XX
PS Claim 2; Fig 5; 99pp; English.

XX This is the nucleotide sequence of full-length cDNA corresponding to a
CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
CC development of platelets and the pathogenesis of leukemias, both in
CC general and in particular those involving the megakaryocytic lineage. The
CC SH3D1A gene maps to the small candidate region for low platelets on
CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
CC AAZ34570-74) suggests that at least 3 isoforms exist. The invention
CC provides methods for the diagnosis and treatment of megakaryocytic
CC abnormality, myeloproliferative disorder, platelet disorder, acute
CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on
CC chromosome 21, low platelets in deletion for 21, association of gains in
CC chromosome 21 with leukemias, neural abnormalities, dysfunctions and
CC disorders including brain malformations and corresponding cognitive
CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
CC also provided for: suppressing cells unable to regulate themselves;
CC screening for a somatic alteration in the SH3D1A gene; monitoring the
CC progress and adequacy of a treatment; monitoring tumour risk progress or
CC megakaryocytic abnormality, myeloproliferative disorder, haematopoietic
CC disorder, platelet disorder or leukaemia; and treatment of a subject
CC (including a prenatal subject) having megakaryocytic abnormality,
CC myeloproliferative disorder, platelet disorder, leukaemia or neural
CC disorder using a nucleic acid that expresses SH3D1A or its antisense
CC nucleic acid

XX
SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.92e-303 Length: 5199

Score: 5895.00 Matches: 1143

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 2 Gaps: 0

US-09-720-934-2 (1-1143) x AAZ34570 (1-5199)

Qy	1	MetAlaGlnPheProThrPheGlySerLeuAspIleTrpAlaIleThrValGlu	20
Db	208	ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCTGGATATCTGGGCCATAAATCTAGAG	267
Qy	21	GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr	40
Db	268	GAAGAGCGAAGACGATGATCAGCAGTTCCATAGTTTAAAGCCAAATATCTGGATTCAATACT	327
Qy	41	GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln	60
Db	328	CGTGATCAAGCTAGTAAACTTTTTCCTCAATCTGGTTACTCAACCTGTTTACGACAG	387
Qy	61	IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle	80
Db	388	ATATGGGCACCTAGCTGACATGAATATATGATGAAGATGGATCAAGTGGAGTTTTCATA	447
Qy	81	AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro	100
Db	448	GCTATGAACACTTATCAAACTGAAGCTACAGGATATACAGTACCCCTCTGCATTCCTCCCT	507
Qy	101	ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle	120
Db	508	GTCATGAACACCAACCACTTTCTTCTAGCGCCAGCATTTGGTATGGAGGTATC	567
Qy	121	AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal	140
Db	568	GCCAGCATGCCACCGCTTACAGCTGTGCTCCAGTGGCAATGGGATCCATTCCAGTTGTT	627

Qy	141	GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn	160
Db	628	GGATGTCTCCACACCTTAGTATCTTGTGTCCACAGCAGCTGTGCCCCCTGGCTAAC	687
Qy	161	GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro	180
Db	688	GGGGCTCCCGCTGTATACAACTCTGCTGCAATTTGCTCATCTGCAGCCACATATGCCA	747
Qy	181	LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys	200
Db	748	AAGAGTTCTTCTTTTAGTAGATCTGGTCCAGGGTCACTAAACACATAATTTACAAAG	807
Qy	201	AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer	220
Db	808	GCACAGTCATTTGATGTGGCCAGTGTCCACACAGTGGCAGAGTGGGTGTTCTCTAGTCA	867
Qy	221	SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu	240
Db	868	TCAAGACTGAATACAGGCAATTTATTCAATAGTCATGACAAACTATGAGTGGACACTTA	927
Qy	241	ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla	260
Db	928	ACAGGTCCTCCCAAGCAAGAACTATTTCTTATGAGTCAAGTTTACCACAGGCTCAGTGGCT	987
Qy	261	SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle	280
Db	988	TCAATATGGAATCTTTCTGACATTTGATCAAGATGGAAACTTACAGCAGAGGAATTTATC	1047
Qy	281	LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro	300
Db	1048	CTGGCAATGCACCTCATTGATGTAGTATGTGTGGCCAACTGTCACCTGCTCTCTGCT	1107
Qy	301	ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle	320
Db	1108	CCAGATACACTTCCACCTTCTTTTAGAAGAGTTCGATCTGGCAGTGGTATATCTGTCTATA	1167
Qy	321	SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln	340
Db	1168	AGCTCAACATCTGTAGATCAGAGGCTACCAGAGGAAACGAGTTTGTAGAAGATGAACAA	1227
Qy	341	GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg	360
Db	1228	CAATTAGAAAAGAAATTTACCTGTAACTTTGAAGATGAAGCCGGAGAACTTTTGAACGT	1287
Qy	361	GlyAsnLeuGluLeuLysArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln	380
Db	1288	GGCAACTGTGAACTGGAGAAACGAGGCAAGCTCTCTCTGGACAGCAGCGCAGAGGAG	1347
Qy	381	GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu	400
Db	1348	GAGCGCTGGCCAGCTGGAGCGCGGAGCAGGAGAGGAGGAGCGTGGAGCGCAGGAG	1407
Qy	401	GlnGluArgLysArgGlnLeuGluLysGlnLeuLysGlnLeuArgGluLeuGlu	420
Db	1408	CAAGAGCGCAAAAGACAACTGGAACTGGAGAAGCAACTGGAAAAAGCAGCGGAGCTAGAA	1467
Qy	421	ArgGlnArgGluGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu	440
Db	1468	CGGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1527
Qy	441	LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGln	460
Db	1528	CTTGAAAGGCAACCACTTGAAGTGGGAAACGGAATCGAAGCGCAAGAACTACTAATCAA	1587
Qy	461	ArgAsnLysGluGlnGluAspIleValIleLeuLysAlaLysLysLysThrLeuGluPhe	480
Db	1588	AGAAACAAAGAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1647
Qy	481	GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg	500
Db	1648	GAATTAGAGCTCTAAATGATAAAAGCATCACTAGAGGAGGAACTTCAAGATATCAGA	1707

lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 239..3886
XX /*tag= a

XX WO9953062-A2.

XX 21-OCT-1999.

XX 16-APR-1999; 99WO-US008371.

XX 16-APR-1998; 98US-0082007P.

XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.

XX Korenberg JR, Chen X;

XX WPT; 1999-633829/54.

XX P-PSDB; AAY32156.

PT Nucleic acid from the human SH3D1A gene and its products, useful for the diagnosis and treatment of myeloproliferative disorders and leukemia.

XX Claim 2; Fig 10; 99pp; English.

XX This is the nucleotide sequence of full-length cDNA (clone 11) corresponding to a novel human SH3 gene, termed the SH3D1A gene, that contributes to the development of platelets and the pathogenesis of leukemias, both in general and in particular those involving the megakaryocytic lineage. The SH3D1A gene maps to the small candidate region for low platelets on chromosome 21. Sequencing of 5 different sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at least 3 isoforms exist. The invention provides methods for the diagnosis and treatment of megakaryocytic abnormality, myeloproliferative disorder, platelet disorder, acute leukemia, neural disorders, thrombocytopenia, platelet disorder on chromosome 21, low platelets in deletion for 21, association of gains in chromosome 21 with leukemias, neural abnormalities, dysfunctions and disorders including brain malformations and corresponding cognitive dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are also provided for: suppressing cells unable to regulate themselves; screening for a somatic alteration in the SH3D1A gene; monitoring the progress and adequacy of a treatment; monitoring tumour risk progress or megakaryocytic abnormality, myeloproliferative disorder, haematopoietic disorder, platelet disorder or leukemia; and treatment of a subject (including a prenatal subject) having megakaryocytic abnormality, myeloproliferative disorder, platelet disorder, leukemia or neural disorder using a nucleic acid that expresses SH3D1A or its antisense nucleic acid

XX Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,97e-301 Length: 5195
Score: 5849.50 Matches: 1143
Percent Similarity: 94.15% Conservative: 0
Best Local Similarity: 94.15% Mismatches: 0
Query Match: 99.23% Indels: 71
DB: 2 Gaps: 1

US-09-720-934-2 (1-1143) x AAZ34572 (1-5195)

Qy 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrrAlaIleThrValGlu 20

Db 239 ATGGCTCAGTTTCCAAACCTTTTGGTCGACGCTGGATATCTGGGCCATAACTGTAG 298

Qy 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40

Db 299 GAAGAGCGAAGCATCATCAGCAGTTCCTAGATTAAAGCCATATCTGGATTCATTACT 358

Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60

Db 359 GGTGATCAAGCTAGAACTTTTTCATCTGGTTACCTCAACCTGTTTATGACAG 418
Qy 61 IleTrrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 419 ATATGGGCACCTAGCTGACATGAATAATGATGAAGAATGGATCAAGTGGAGTTTCCATA 478
Qy 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTrrGlnLeuProSerAlaLeuProPro 100
Db 479 GCTATGAAACTTATCAAACTGAAGCTCAAGGATATCAGCTACCTCTGCACTTCCCT 538
Qy 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyIle 120
Db 539 GTCATGAACAGCAACACAGTTGCTATTTCTAGCGCACCCAGCATTTGGTATGGAGGTATC 598
Qy 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db 599 GCCAGCATGCCACCGCTTACAGCTGTTGCTCCAGTGCATATGGGATCCATTCACATTGTT 658
Qy 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db 659 GGAATGCTCCAACTAGTATCTTCTGTTCCACAGCAGCTGTGCCCCCTGGCTAAC 718
Qy 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db 719 GGGGCTCCCTGTTATACAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
Qy 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db 779 AAGAGTTCTTCTCTTTAGTAGATCTGGTCCAGGGTCACAACTAAACACATAATACAAAG 838
Qy 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrrAlaValProGlnSer 220
Db 839 GCACAGTCTTTGATGTGGCAGTGTCCACAGTGGCAGAGTGGGCTGTTCTCTCAGTCA 898
Qy 221 SerArgLeuLysTrrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db 899 TCAAGACTGAAATCAGGCAATTTTCAATAGTCATGACAAACTATGAGTGGACACTTA 958
Qy 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
Db 959 ACAGTCTCCCAAGCAAGAACTATTCTTATGAGTCAAGTTTACCAAGCTGAGTGGCT 1018
Qy 261 SerIleTrrAsnLeuSerAspIleAspGlnLysLeuThrAlaGluGluPheIle 280
Db 1019 TCAATATGGAATCTTCTGACATTCATGATCAAGATGGAAACTTACAGCAGAGGAATTTATC 1078
Qy 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
Db 1079 CTGGCAATGCACCTCATTGATGATGATGCTGCGCCAAACCACTGCCACCTGCTGCTGCT 1138
Qy 301 ProGluTrrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
Db 1139 CAGAATACATTCACCTTCTTTTAGAAGATTCGATCTGGCAGGGTATATCTCTCATATA 1198
Qy 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
Db 1199 AGCTCAACATCTGTAGATCAGAGGCTTACAGAGGAACCACTTTTAGAAGATGAACACAA 1258
Qy 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
Db 1259 CAATTAGAAAAGAAATTAACCTGTACCTGTTAAGATAAGAGCGGAGAACTTTTGAACGT 1318
Qy 361 GlyAsnLeuGluLeuGluLysArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
Db 1319 GCGAACCTTGGAACTGGAAACGAAGCAAGCTCTCTTGGAAACAGCAGCGCAGAGGAGCAG 1378
Qy 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluGlnGlu 400
Db 1379 GAGCGCTGGCCAGCTGGAGCGGCGGAGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1438
Qy 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420

Db 1439 CAAGAGCGCAAAAGACAACTGGAACTGGAGAGCAACTGGAAAGACGCGGGAGCTAGAA 1498
QY 421 ArgGlnArgGluGluArgArgLysGluileGluArgArgGluAlaLysArgGlu 440
Db 1499 CGGACAGAGAGGAGGAGAGGAGGAGAAATTTGAGAGGCGAGAGGCTGCAAAACGGAA 1558
QY 441 LeuGluArgGlnArgGlnLeuGluTTPGluArgAsnArgArgGlnLeuLeuAsnGln 460
Db 1559 CTTGAAAGGCAACGACAACTTGAGTGGGAAACGGAATCGAAGGCAAGAACTACTAAATCAA 1618
QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
Db 1619 AGAAACAAAGACACAGAGGACNTAGTTGTACTGAACCAAGAAAGAACTTTGGAAATTT 1678
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
Db 1679 GAATTTAGAACTCTAAATGATAAAAGCATCACTAGAGGGAACCTTCAAGATATCAGA 1738
QY 501 CysArgLeuThrThrGlnArgGlnGluileGluSerThrAsnLysSerArgGluLeuArg 520
Db 1739 TGTGCAATTGACCCCAAGAGCAAGAAATTTGAGAGCAACAACAAATCTAGAGAGTTGAGA 1798
QY 521 IleAlaGluileThrHisLeuGlnGlnGlnLeuGlnSerGlnGlnMetLeuGlyArg 540
Db 1799 ATTGCCGAATCACCCATCTACAGCAACAATTTACAGGAATCTCAGCAAAATGCTTTGGAAGA 1858
QY 541 LeuileProGluLysGlnileLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1859 CTTATTCCGAAAAACAGATACTCAATGACCAATTTAAACCAAGTTTCAGCAGAAACAGTTTG 1918
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1919 CACAGAGATTCACTGTGTACACTTAAAGAGCCCTTAGAAGCAAAAGAACTAGTTCGCGAG 1978
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluile 600
Db 1979 CACTACGAGACCACCTGGATGAAGTGGAGAAAGAACTAGATCAAACTACAGGAGATT 2038
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluileHisAsnLysGlnGlnLeuGln 620
Db 2039 GATATTTTCAATAATCAGCTGAAGGAACTAAGAGAAATACACAATAAGCAACAACCTCCAG 2098
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnArgLysIleIle 640
Db 2099 AAGCAAAAGTCCATGGAGGCTGAACCACTGAAACAGAAAGAAACAGAAAGATCATA 2158
QY 641 GluLeuGluLysGlnLysGluGluAlaGlnArgAlaGlnGluArgAspLysGlnTTP 660
Db 2159 GAATTAGAAAAACAAAAAGAGAGCCCAAGACGAGCTCAGGAAAGGGACAGCAGTGG 2218
QY 661 LeuGluHisValGlnGlnAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 2219 CTGGAGCATGTGCAGCAGGAGGAGCAGCATCAGAGACCAAGAAAACTCCACGAAGAGGAA 2278
QY 681 LysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLysGlnGlnGlu 700
Db 2279 AAACGAAAGGGAGGAGAGTGTCAAAAAAGAGGATGGCGAGAAAAAGCAACAGGAA 2338
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnProAlaLysProAlaVal 720
Db 2339 GCACAAGACAGCTGGGTGGCTTTTCCATCAACACCAAGAACACGCTAAGCCAGCTGTC 2398
QY 721 GlnAlaProTTPSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGlnAsnVal 740
Db 2399 CAGGCACCCCTGGTCCACTGCAGAAAAAGGTCCACTTACCATTCTGCACAGGAAAAATGA 2458
QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluileThr 760
Db 2459 AAAGTGGTGTATTACCGGCACTGTACCCCTTTGAAATCCAGAGCCCATGATGAAATCACT 2518
QY 761 IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTTPLeu 780
Db 2519 ATCCACCCAGGAGACATAGTATGTTGGATGAAAGCCAAACTGGAGAACCCCGCTGGCTT 2578

QY 781 GlyGlyGluLeuLysGlyLysThrGlyTTPPheProAlaAsnTyrAlaGluLysIlePro 800
Db 2579 GGAGGAGAAATTAAGAGAAAGACAGGGTGGTCCCTGCAAACTATGCAGAGAAATCCCA 2638
QY 801 GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro 820
Db 2639 GAAAATGAGGTTCCTCGCTCCAGTGAAACAGTCACTGATTCAACATCTGCCCTTCCGCC 2698
QY 821 LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr 840
Db 2699 AAATGGCTTGGCTGAGACCCCGCCCTTTGGCAGTAACTCTTCCAGAGCCCTCCACG 2758
QY 841 ThrProAsnAsnTTPAlaAspPheSerSerThrTTPProThrSerThrAsnGluLysPro 860
Db 2759 ACCCTTAATAACTGGGCGGACTTCAGTCCACGTGGCCCAACAGCAGATGAGAAACCA 2818
QY 861 GluThrAspAsnTTPAspAlaTTPAlaAlaGlnProSerLeuThrValProSerAlaGly 880
Db 2819 GAAACGGATAACTGGGATGCATGGGAGCCCGCCCTCTCTCACCGTTCCAAAGTCCCGGC 2878
QY 881 GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro 900
Db 2879 CAGTTAAGGAGAGGTCGCGCTTTACTCCAGCCACGCGCACTGGCTCCTCCCGCTCTCT 2938
QY 901 ValLeuGlyGlnGlyGluLysValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArg 920
Db 2939 GTCTAGGCCAGGTTGAAAAGGTGGAGGGGTACAAGCTCAAGCCCTATATCTCTGGAGA 2998
QY 921 AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGln 940
Db 2999 GCCAAAAAGACAAACCACTTAAATTTTACAAAAATGATGTCATCACCGCTCTGGAACAG 3058
QY 941 GlnAspMetTTPThrPheGlyGluValGlnGlyGlnLysGlyTTPPheProLysSerTyr 960
Db 3059 CAAGCATGTGGTGGTTGGAGAAAGTTCAAGGTTCAGAGGGTTGGTTCCCAAGTCTTAC 3118
QY 961 ValLysLeuIleSerGlyProIleArgLysSerThrSerMetAspSerGlySerSerGlu 980
Db 3119 GTGAAACTCATTTTCAGGGCCCAATAAGGAAGTCTACAAGCATGGATTCTGGTCTTCAGAG 3178
QY 981 SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu 1000
Db 3179 AGTCCTGCTAGTCTAAAGCGAGTAGCTCTCCAGCAGCAAGCCGCTCGTTTCGGGAGAA 3238
QY 1000 ----- 1000
Db 3239 GAATTTATTGCCATGTACACTTACGAGAGTTCTGACGAAGGAGATTAACTTTTCAGCAA 3298
QY 1000 ----- 1000
Db 3299 GGGGATGTGATTTTGGTTTACCAAGAAAGATGGTGACTGGTGGACAGGAACAGTGGCGAC 3358
QY 1000 ----- 1000
Db 3359 AAGGCCGAGTCTTCCCTTTCTAACTATGTGAGGCTTTAAAGATTCAGAGGGCTCTCGAACT 3418
QY 1001 ----- 1009
Db 3419 GCTGGGAAAAACAGGGAGTTTAGAAAAAAACCTGAAATTTGCCAGGTTATTTGCTCATAC 3478
QY 1010 ThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuIleLeuArgLys 1029
Db 3479 ACCGCCACCGCCCGAGCAGCTCACTCTCGCCCTGGTTCAGTCACTGATTTTATCCGAAA 3538
QY 1030 LysAsnProGlyGlyTTPThrGluGlyGluLeuGlnAlaArgGlyLysLysArgGlnIle 1049
Db 3539 AAGAACCCAGGTGGATGGTGGGAAGGAGAGTGCAGACACGTGGGAAAAAGGCCAGATA 3598
QY 1050 GlyTTPPheProAlaAsnTyrValLysLeuLeuSerProGlyThrSerLysIleThrPro 1069
Db 3599 GGCTGTTTCCCACTAAATTTATGTAAGCTTCTAAGCCCTTGGGACGAGCAAAATCACTCCA 3658

QY 261 SerIleThrAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
Db 1047 TCAATATGGAAATCTTCTGACATTTGATCAAGATGGAAATCTTACAGCAGAGGAATTTATC 1106
QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
Db 1107 CTGGCAATGCACCTCATTTGATGTAGCTATGTCTGGCCAAACCCTGCCACCTGTCTGCCT 1166
QY 301 ProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle 320
Db 1167 CCAGAAATACATTCACCTCTCTTTAGAGAGTTCCGATCTGGCAGTGTATATCTGTCTATA 1226
QY 321 SerSerThrSerValAspGlnArgLeuProGluCluProValLeuGluAspGluGlnGln 340
Db 1227 AGCTCAACATCTGTAGATCAGAGCTACAGAGAACCTAGTTTAGAAGATGAACAA 1286
QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
Db 1287 CAATTAGAAAAGAAATTTACCTGTAACTTTGAAGATGAAGCGGGAACCTTTGAACGT 1346
QY 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnArgLysGluGln 380
Db 1347 GGCACCTTGAATCGAGAGAACGAAAGCAAGCTCTCTCGAAACAGCAGCGCAAGGAGCAG 1406
QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu 400
Db 1407 GAGCGCTTGGCCAGCTGGAGCGGCGGAGAGAGAGAGAGCGGTGAGCGCCAGGAG 1466
QY 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
Db 1467 CAAGAGCGCAAAAGACAACTTGGAACTGGAGAGCACTGGAAAGAGCGGAGCTAGAA 1526
QY 421 ArgGlnArgGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu 440
Db 1527 CGCAGAGAGAGGAGAGAGAGAGAGAAATTTGAGAGCGAGAGGCTGCAAAACGGGAA 1586
QY 441 LeuGluArgGlnArgGlnLeuGluTyrGluArgAsnArgArgGlnGluLeuLeuAsnGln 460
Db 1587 CTTGAAGGCAACGCAACTTGAATGGAGGAAACGGAATCGAAGGCAAGAACTACTAAATCAA 1646
QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
Db 1647 AGAAACAAAGAAACAAGAGGACATAGTTGACTGAAGCAAGAAAGAACTTTGGAAATTT 1706
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
Db 1707 GAATTTAGAGCTCTAAATGATAAAAGCATCAACTAGAGGGGAACTTCAAGATATCAGA 1766
QY 501 CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg 520
Db 1767 TGTGGAATTTGACCCCAAGGCAAGAAATTTGAGAGCACAAACAAATCTAGAGAGTTGAGA 1826
QY 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGlnGlnMetLeuGluArg 540
Db 1827 ATTGCGGAATCAACCATCTACAGCAACAATTACAGAACTCTCAGCAAAATGCTTGGAGA 1886
QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1887 CTTATTTCCAGAAAACAGATACTCAATGACCAATTAACAACTTCAAGTTCAGCAGAACAGTTTG 1946
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1947 CACAGAGATTCATTTGACCTTAAAGAGCGCTTAGAAGCAAAAGAACTAGTCTGGCAG 2006
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
Db 2007 CACCTACGAGACCAACTGGATGAAGTGAGAGAAAGAACTAGATCAAAACTTACAGGAGATT 2066
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
Db 2067 GATATTTTCAATTAATCAGCTGGAAGGAACTAAGAGAAATACACAATAAGCAACAACCTCCAG 2126
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnArgLysIleIle 640

Db 2127 AAGCAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAAAGAACCAAGATCAT 2186
QY 641 GluLeuGluLysGlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTyr 660
Db 2187 GAATTAGAAAAACAAG 2246
QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 2247 CTGGAGCATGTGAGCAGGAGGAGCAGCATCAGAGACCAAGAAACTCCACGAGAGGAA 2306
QY 681 LysLeuLysArgGluGluSerValLysLysAspGlyGluGluLysGlyLysGlnGlu 700
Db 2307 AAATGAAAAGGAGAGAGAGTGTCAAAAGAGAGAGTGGCAGGAAAGGCAACAGGAA 2366
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal 720
Db 2367 GCACAGACAACTGGGTGGCTTTTCCATCAACACCAAGACAGCTAGCAGCTGTC 2426
QY 721 GlnAlaProThrSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
Db 2427 CAGGCAACCTGGTCCACTGTCAGAAAAAGGTCCACTTACCATTTCTGCACAGGAAATGTA 2486
QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
Db 2487 AAAGTGTGTATTACCGGGCAGCTGTACCCCTTTGAAATCCAGAGCCATGTATGAATCACT 2546
QY 761 IleGlnProGlyAspIleValMet-----ValAspGluSerGlnThrGly 775
Db 2547 ATCCAGCAGGAGACATAGTCATGTTAAAGGGGAATGGGTGATGAAAGCCAACTGGA 2606
QY 776 GluProGlyTyrLeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyr 795
Db 2607 GAACCGCGCTGGCTTGGAGAGAAATTAAGAGAGAGACAGGGTGGTTCCCTGCAAACTAT 2666
QY 796 AlaGluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThr 815
Db 2667 GCAGAGAAATCCAGAAATAGGTTCCCGCTCCAGTGAACCAAGTGAATTCACCA 2726
QY 816 SerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSer 835
Db 2727 TCTGCCCTTGGCCCAAACTGGCTTGGCTGAGACCCCGCCCTTTGGAGAGTAACCTCT 2786
QY 836 SerGluProSerThrThrProAsnAsnThrAlaAspPheSerSerThrThrProThrSer 855
Db 2787 TCAGAGCCCTCCACGACCCCTAATAACTGGCGGCGACTTCAGCTCCACGTCGCCACGAGC 2846
QY 856 ThrAsnGluLysProGluThrAspAsnThrAspAlaTyrAlaTyrAlaGlnProSerLeuThr 875
Db 2847 ACCAATGAGAAACCAAGAACGGATTAATGGGATGTCATGGGCGCCCGCCCTCTCTCACC 2906
QY 876 ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly 895
Db 2907 GTTCCAAGTGGCGCCAGTTAAGCAGAGAGTCCGCTTTACTCCAGCCAGGCCACTGGC 2966
QY 896 SerSerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAla 915
Db 2967 TCCTCCCGCTCTCTCTGTCTAGGCCAGGTTGAAAGGTGGAGGGGCTACAAGCTCAAGCC 3026
QY 916 LeuTyrProThrArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIle 935
Db 3027 CTATATCTTGGAGAGCCAAAAGAACACCACTTAAATTTTAAACAAAATGATGTATC 3086
QY 936 ThrValLeuGluGlnGlnAspMetTyrTyrPheGlyGluValGlnGlnGlyLysGlyTyr 955
Db 3087 ACCGTCTTGGACAGACAGACATGTGTGGTTGGAGAGTTCAAGGTTCAGAGGGTGG 3146
QY 956 PheProLysSerTyrValLysLeuIleSerGlyProIleArgLysSerThrSerMetAsp 975
Db 3147 TTCCCAAGTCTTACGTGAACCTCAATTCAGGGCCCATTAAGGAAGTCTACAAGCATGGAT 3206
QY 976 SerGlySerSerLeuSerProAlaSerLeuLysArgValAlaSerProAlaLysPro 995


```
Db 3207 TCTGGTCTTTCAGAGTCTCTAGTCTAAAGCGAGTAGCTCTCCAGCAGCCAGCCG 3266
QY 996 ValValSerGlyGlu----- 1000
Db 3267 GTCGTTTCGGGAGAGAATTTATTGTCATGTACACTTTACGAGAGTTCTGAGCAAGAGAT 3326
QY 1000 ----- 1000
Db 3327 TTAACCTTTTCAGCAAGGGATGTGATTTTGGTTTACCAGAAAGATGGTACTGGTGACA 3386
QY 1000 ----- 1000
Db 3387 GGAACAGTGGCGACAGCGGAGTCTTCCCTTCTAACTATGTGAGCTTTAAAGATTCA 3446
QY 1001 -----GluLeuAlaGln 1004
Db 3447 GAGGGCTCTGAACTGCTGGGAAACAGGGAGTTTAGGAAAAAACCTGAAATTTGCCAG 3506
QY 1005 ValLeuAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeu 1024
Db 3507 GTTATGCTCATACCGCCAGCCGCGCCGAGCAGCTCACTCTGCCCCCTGGTCAGCTG 3566
QY 1025 IleLeuIleArgLysLysAsnProGlyGlyTyrTrpGluGlyGluLeuGlnAlaArgGly 1044
Db 3567 ATTTTGATCCGAAAAAGAACCCAGGTGGATGGTGGGAAAGAGAGCTGCAAGACAGCTGGG 3626
QY 1045 LysLysArgGlnIleGlyTyrPheProAlaSerTyrValLysLeuLeuSerProGlyThr 1064
Db 3627 AAAAAGCGCCAGATAGGCTGTTCCAGCTAATTATGTAAAGCTTCTAAAGCCCTGGGACG 3686
QY 1065 SerLysIleThrProThrGluProProLysSerThrAlaLeuAlaAlaValCysGlnVal 1084
Db 3687 AGCAAAATCATCTCCACAGACCCACTAAGTCAACAGATTAGCGGCAAGTGCACAGGTG 3746
QY 1085 IleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLysGlyGln 1104
Db 3747 ATTGGGATGTACGACTACACCGCGCAGAAATCAGCATGAGCTGGCCCTTCAACAAGGCCAG 3806
QY 1105 IleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGluValAsnGlyGln 1124
Db 3807 ATCATCAACGCTCTCAACAAGGAGGACCTGACTGCTGGTGGAAAGGAGAAAGTCAATGGCAA 3866
QY 1125 ValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
Db 3867 GTGGGGCTCTTCCCATCCAAATTATGTGAAGCTGACCACAGCATGACCCAGCCAGCCAG 3923
RESULT 4
ID AAS84763
XX AAS84763 standard; cDNA; 7435 BP.
AC AAS84763;
XX
XX
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #20567.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
```

```
XX WPI; 2001-639362/73.
DR P-PSDB; ABG20576.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 20567; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 3.5e-279 Length: 7435
Score: 5441.00 Matches: 1122
Percent Similarity: 90.67% Conservative: 5
Best Local Similarity: 90.27% Mismatches: 16
Query Match: 92.30% Indels: 100
DB: 5 Gaps: 4
```

```
US-09-720-934-2 (1-1143) x AAS84763 (1-7435)
QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 263 ATGGCTCAGTTTCCACACCTTTTGGTGGAGCTGGATACTGGGCCATAACTAGAG 322
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 323 GAAAGAGCGAAGCATGATCAGCAGCTCCATAGTTTAAAGCCAATATCTGGATTCAATTACT 382
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 383 GGTGATCAAGCTAGAAACTTTTTTTTCAATCTGGGTACCTCAACCTGTTTTAGCACAG 442
QY 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 443 ATATGGGACACTAGCTGACATGATATGATGGAAGATGATCAAGTGGAGTTTTCCATA 502
QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 503 GCTATGAAACTTATCAAACTGAAGCTACAAAGGATATCAGTACCTCTGCACTTCCCCCT 562
QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db 563 GTCATGAAACAGCAACCACTGCTATTTCTACGCCACACAGCAATTTGGTATGGAGGTATC 622
QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIle-ProValVa 140
Db 623 GCCAGCATGCCACCGCTTACAGCTGTGTCTCCAGTGGCAATGGGATCCATTCCCAAGTTGT 682
```


Db 2843 CTCACGTCGCCACACGACGACGATGAGAAACGAGAAACGATACCTGCGATGCGGC 2902

Qy 869 alaGln-ProSerLeuThrValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheT 889

Db 2903 AGCCAGGCGCTCTCTCACCGTTCCAGTCCGCGCAGTTAAGCGACAGGTCGCGCTTTA 2962

Qy 889 hrProAlaThrAlaThrGlySerProSerProValLeuGlyGlnGlyValG 909

Db 2963 CTCACGACCGCCACTGGCTCTCCCGCTCTCTGTGCTAGCCAGGTCGAAAGGTGG 3022

Qy 909 luGlyLeuGlnAlaGlnAlaLeuTyProTrpArgAla-lysLysAspAsnHisLeuAsn 928

Db 3023 AGGGCTTACAGCTCAAGCCCTATATCTTTGGAGAGCCAAAAAGACCAACCACTTAAAT 3082

Qy 929 PheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGlyGlu 948

Db 3083 TTTTACAAAAATGATGTCATCCGCTCTCGAAACAGCAACATGTGTGTGGAGAA 3142

Qy 949 ValGlnGlyGlnLysGlyTrpPheProLysSerTyrValLysLeuIleSerGlyProIle 968

Db 3143 GTTCAAGGTGAGAGGGTGTGTTCCCAAGTCTTACGTGAACATCTATTTAGGGCCCAT 3202

Qy 969 ArgLysSerThrSerMet-AspSerGlySerSerGluSerProAlaSerLeuLysArgVa 988

Db 3203 AGGAAGTCTCAAGCATGGGATCTGGTTCTTCAGAGAGTCTCTGCTAGTCTAAAGCGAGT 3262

Qy 988 lAlaSerProAlaLysProValValSerGlyGlu----- 1000

Db 3263 AGCTCTCCAGCAGCAAGCGGTGTTTCGGGAGAGGAATTTATGTCATTTACACTT 3322

Qy 1000 ----- 1000

Db 3323 ACAGAGTCTTACGCAAGGAGATTTAACTTTTACGCCAGGGGGATGTGATTGGT 3382

Qy 1000 ----- 1000

Db 3383 ACCAAGAAGATGGGACTTGTGGACAGAACAAAGTGGCGACAGCGCGAGTCTTCCC 3442

Qy 1000 ----- 1000

Db 3443 TTCTAACTATGTAGGCTTAAAGATTCAGAGGGCTCTGGAACCTTCGGAAACAGGAG 3502

Qy 1001 -----GluIleAla-GlnValIleAlaSerTyrThrAlaThrGlyProG 1015

Db 3503 TTTAGGAAAAAACCTGAATTTGCCCGAGTTATGCTTCATACCGCCACCGCCCG 3562

Qy 1015 luGlnLeuThrLeuAlaProGlyGlnLeuIleAlaGlyLysAsn-ProGlyGly 1034

Db 3563 AGCAGTCACTCTGCCCCCTGTGCTAGCTGATTTGATCCGAAAAAGAACCCCGGTGA 3622

Qy 1035 TrpTrpGluGlyGluLeuGlnAlaArgGlyLysArgGlnIleGlyTrpPheProAla 1054

Db 3623 TGGTGGAGAGAGAGTGAACGACGTCGTGGGAAAAAGCGCCAGATGGCTTGGTCCAGT 3682

Qy 1055 AsnTyrValLysLeuLeuSerProGlyThrSerLysIleThrProThrGluProLys 1074

Db 3683 AATTATGTAAGCTTTAAACCTGGAGCGAGCAAAATCACTCCACAGACCACTAAG 3742

Qy 1075 SerThrAlaLeuAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsn 1094

Db 3743 TCAACAGCATTAGCGGCGAGTGTGCCAGGTGATTTGGGATGTACGACTACACCGCGCAAT 3802

Qy 1095 AspAspGluLeuAlaPheAsnLys-GlyGlnIleIleAsnValLeuAsnLysGluAspPr 1114

Db 3803 GACGATGAGTGGCTTCAACAGGCGGCGAGATCATCAAGCTCTCAACAGAGGAGCCC 3862

Qy 1114 oAspTrpTrpLysGlyGluValAsnGlyGlnValGlyLeuPheProSerAsnTyrVally 1134

Db 3863 TGACTGTGGAAGAGGAGAGTCAATGGACAGTGGGGCTCTTCCCATCCCAATATGTGAA 3922

Qy 1134 sLeuThrThrAspMet-AspProSerGln 1143

Db 3923 GCTGACCACAGACATGGAGCCCAAGCCAG 3951

RESULT 5

AAZ39009
ID AAZ39009 standard; cDNA; 3723 BP.

XX AAZ39009;

XX 28-FEB-2000 (first entry)

XX Mouse Ees1 coding sequence.

XX Mouse; murine; Ees1; Ees2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EF-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.

OS Mus sp.

XX W09955728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 98CA-02230201.

XX 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

XX Egan SE, Wang W, Sengar A;

XX WPI; 2000-052802/04.

XX P-PSDB; AAY57444.

XX New nucleic acid encoding Ees1 and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.

XX Claim 6; Page 40-42; 99pp; English.

XX The present sequence encodes mouse Ees1. The present invention
XX specifically describes mammalian Ees1 and 2 proteins (I) and their splice
XX variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
XX are involved in regulation of clathrin-mediated endocytosis (as a complex
XX with Ees15 protein), vesicular trafficking and actin cytoskeleton.
XX Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
XX mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
XX polynucleotide; agents that downregulate expression of Ees1 genes or
XX antagonists of an Ees1 binding partner are used to treat diseases
XX associated with undesirable endocytosis and resulting changes in cellular
XX function. Particularly overexpression of Ees1 is used to block clathrin-
XX mediated endocytosis in vivo or in cell cultures, while administration of
XX (I) is used to promote endocytosis of selected cells. (Ant)agonists of
XX (I) or Ab are used to suppress abnormal proliferation of cells that can
XX be stimulated to proliferate by a growth factor receptor; and similar
XX compounds (also inactive Ees1 mutants) can be used to prevent viral
XX infection. Endocytosis may also be regulated, in vivo or in cell
XX cultures, by forming an Ees1-Ees15 complex, then binding dynamitin to the
XX complex. Generally conditions that can be treated include cancer;
XX abnormal cell division or migration; viral infection; or abnormal
XX receptor signalling, tissue development or synaptic transmission

XX Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,49e-279 Length: 3723
Score: 5437.50 Matches: 1059
Percent Similarity: 90.28% Conservatives: 37
Best Local Similarity: 87.23% Mismatches: 45
Query Match: 92.24% Indels: 73
DB: Gaps: 3

US-09-720-934-2 (1-1143) x AAZ39009 (1-3723)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
DB 1 ATGGCTCAGTTTCCACACCTTTGGTGGTAGGCTGGATGCTGGGCCATAAATGTGGAG 60
QY 21 GluArgAlaIlyHisAspGlnGlnPheHisSerLeuIysProIleSerGlyPheIleThr 40
DB 61 GAAAGGGCCAGCATGACCGACAGTTCCTTAGGCTGAAGCCGATAGCGGATTTATTACT 120
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
DB 121 CGTGATCAAGCGAGGAATTTTTTTTCCCAATCTGGGTACCTCAGCCTGCTTAGCACAA 180
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
DB 181 ATATGGGCGCTAGCGGACATGAATAACGATGGAAGGATGATCAAGTGGAAATTTCCATA 240
QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
DB 241 GCCATGAAGCTTATCAAACTGAAGCTACAAGGATATCAGTCCCTCCACACTTCCCCCT 300
QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
DB 301 GTCATGAAACAGCAACAGTGGCTATTTCCAGTGCACCACTTTGGTATAGGAGGAT 360
QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
DB 361 GCTAGCATGCCACCACTCAGCTGTTGCTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 420
QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
DB 421 GGNATGTCACCCCTTAGTATCTTCTGCTCCACAGCAGCTGCTCCCTGGCTTAAC 480
QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
DB 481 GGCGCTCCTCCGCTCATACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
DB 541 AAGAGTTCCTCTTCCAGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGlnuTrpAlaValProGlnSer 220
DB 601 GCACATCATCTCAGTGTCCGACGCGCCCTCCAGCAGCAATGGCTGTCCTCAGTCA 660
QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
DB 661 TCAAGGCTGAATACAGGCAGTTATTCAACAGCCAGCACAACTATGATGACACTTA 720
QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
DB 721 ACAGGTCCCGCAGGCAAGAACTATTCTCATGCAATCAAGTTTACCCCGAGGCTCAGCTG 780
QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
DB 781 TCATATGAATGAAATCTTCTGCATGATCAAGATGGAATACTACTCGCAGAAATTTATC 840
QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
DB 841 CTAGCTATGCACCTAATTGATGTTGCCATGCTGCTGGTCAGCACTGCGCGCCGCTGCCT 900
QY 301 ProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle 320
DB 901 CCAGATATACATCCCTCTCTTCAAGAGATGCTGCTCCGCGAGTGGGATGTCCTGCATA 960
QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
DB 961 AGCTCTCTCTCTGTGATCAGAGGCTGCTCAGAGCCGCTGTCAGAGGATGAGCAGCAG 1020
QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
DB 1021 CCA---GAGAAGAAACTGCTGTCATTTGAAGATAAAGAGCGGAGAACTTCGAGCGA 1077
QY 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnArgLysGluGln 380

DB 1078 GGCAGTGTGAGCTGGAGAAGCCCGCAGCGCTCTTGGAGCAGCAGCGCAAGAGCAG 1137
QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnArgLysGluArgGluArgGlnGlu 400
DB 1138 GAGCGTTGGCTCAGCTGGAGCGCGCGAGCAGAGGAAAGAGCGGAGCGCGAGGAG 1197
QY 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuLysGlnArgGluLeuGlu 420
DB 1198 CAGGAGGCCAAGCGCAGCTGGAGTGGAGAAGCAGCTGGAGAAGCAGCGCGAGCTGGAG 1257
QY 421 ArgGlnArgGluGlnGluArgArgLysGluIleGluArgArgGluAlaIleLysArgGlu 440
DB 1258 CGGCAGCGAGAGGAGGAGGAGGAGGAGATCGAGAGGCGCGAGGCGCGAAACCGGAA 1317
QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgGlnGlnLeuLeuAsnGln 460
DB 1318 CTGGAAAGGCGAGCGCAACTTGAATGGAAACCGAACCGGAGCAGGAACCTCTGAATCAG 1377
QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
DB 1378 AGGAACAAGGAGCAGGAGGCGCACCGTGTCTCTGAGSCAAGGAGGAGACTCTGGAGTTT 1437
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
DB 1438 GAGTTAGAAGCTCTGAATGCAAAAAGCATCAGCTAGAAGGAAAACCTTCAGGATATCAGG 1497
QY 501 CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg 520
DB 1498 TGTCGACTGGCAACCCAGAGGCAAGAAATGAGAGCAGCAACAAAGTCTAGAGAGCTAAGA 1557
QY 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGlnGlnMetLeuGlyArg 540
DB 1558 ATTGTGTAATCAACCCACTTACAGCAGCAGTTCGAGGAATCTCAGCAAAATGCTTGGAGA 1617
QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
DB 1618 CTTATTCCAGAGAAACAGATACTCAGTGACCACTTAAACCAAGTCCAGCAGCAAGCTTG 1677
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
DB 1678 CATAGAGACTCGTCTTCCCTCAAAGAGCCCTTGAAGCAAGAGAGCTGGCCCGCAG 1737
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
DB 1738 CAGCTCCGGAGCAGCTGGACGAGGTGGAGAGAGAGACCAAGGTCAAAGCTGCAGAGATT 1797
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
DB 1798 GATGTTTCAACAACCAACGAGCTGAAGGAACCTGAGAGATATATAGCAAAACAGCACTCCAG 1857
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnArgLysIleIle 640
DB 1858 AAGCAGAGTCTCTGGAGCAGCGCCTGAAACAGAAAGAGCAGGAGAGAGAGAGCCTG 1917
QY 641 GluLeuGluLysGlnLysGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrp 660
DB 1918 GAGTTAGAGAAGCAAAAGGAGAGCGCTCAGAGCAGGATTCAGGAAAGGAGCAAGCAATGG 1977
QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
DB 1978 CTGGACATGTGAGCAGGAG---GAGCAGCAGCCCGCCCGGAAACCCACAGGAGGAGC 2034
QY 681 LysLeuLysArgGluGluSerValLysLysAspGlyGluLysGlyLysGlnGlu 700
DB 2035 AGACTGAAGAGGAGAGCAGTGTCAAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAG 2094
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGlnProAlaLysProAlaVal 720
DB 2095 ATGCAAGCAAGCAGAGTGGCTTTTCCATCCGATCAGAGCAGCAGTAAAGCTGGCCACC 2154
QY 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740

Db 2155 CAGGACCCCTGGTCTACCCAGAGAAAGGCCCTTACCATTCTGCACAGGAGAGTGTA 2214
 QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
 Db 2215 AAAGTGGTATATTACCGAGCGCTGTACCCCTTTGAATCCAGAAGTCACGATGATCACC 2274
 QY 761 IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrTrpLeu 780
 Db 2275 ATCCAGCCAGGAGATATAGTCATGTGGTGAAGCCAGAGCTGAGAGCCAGGATGGCTT 2334
 QY 781 GlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 800
 Db 2335 GAGGAGAGCTGAAGGGAAGACGGGATGTTCCCTGCAAACTATGCAGAAAAGATTCCA 2394
 QY 801 GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro 820
 Db 2395 GAAATGAGGTTCCTCCACTCCAGCCAAACCAAGTCACCGATCTGCATCTGCCCTGCCCCC 2454
 QY 821 LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr 840
 Db 2455 ARACTGGCTCTGGCTGAGACCCCTGCTCTTTCAGTGAGTACCTCTTCTGAGCCCTCCACA 2514
 QY 841 ThrProAsnAsnTrpAlaAspPheSerSerThrTyrProThrSerThrAsnGluLysPro 860
 Db 2515 ACCCCCAACAACCTGGCAGACTTCAGTTCACAGTGGCCAGCAGCTCAAAACGAGAAGCA 2574
 QY 861 GluThrAspAsnTrpAspAlaTrpAlaGlnProSerLeuThrValProSerAlaGly 880
 Db 2575 GAAACGGACAACCTGGGATACGTGGCGGCTCAGCGCTTCTCTGACCGTACCTAGTGGCTGGC 2634
 QY 881 GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro 900
 Db 2635 CAGTTACGGCAGAGATCAGCTTTTACCCACGACAGCCACTGGCTCTCTCCCATCTCCC 2694
 QY 901 ValLeuGlyGlnGlyGluLysValGluLysValGluLysAlaGlnAlaLeuTyrProTyrArg 920
 Db 2695 GTCTTGGCCAGGCTGAAAGGTGGAAGGGCTACAAGCGCAAGCCCTGTATCCCTGGAGA 2754
 QY 921 AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGln 940
 Db 2755 GCCAAAAAGACAACCACTTAAATTTTAAACAAAAGTGACGTCAATCCGCTTCTGGACAG 2814
 QY 941 GlnAspMetTyrTrpPheGlyGluValGlnGlyGlnLysGlyTyrPheProLysSerTyr 960
 Db 2815 CAAGACATGTGTGTGTGGAGAGTTCAAGTCAAGAGGTGTGGTTCCTCCCAAGTCTTAC 2874
 QY 961 ValLysLeuIleSerGlyProIleArgLysSerThrSerMetAspSerGlySerSerGlu 980
 Db 2875 GTGAACTCAITTCAGGGCCGTGAAGGAAATCCACAGCATCGATACTGGGCCCTACTGAA 2934
 QY 981 SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu 1000
 Db 2935 AGTCTGTAGTCTAAAGAGAGTGGCTTCCCGCGCCGCAAGCCAGCCATTCCCGGAGAA 2994
 QY 1000 ----- 1000
 Db 2995 GAGTTATTGCCATGTACATACAGAGATTCAGCAAGGAGATTAACTTTTCAGCAA 3054
 QY 1000 ----- 1000
 Db 3055 GGGGATGTGATTGTGTTTACCAGAAAGATGGTGTGAGCGGAACGGTGGCGGAC 3114
 QY 1000 ----- 1000
 Db 3115 AAGTCCGGAGTCTTCCCTTTCTAACTATGTGAGGCTTAAAGATTCAGAGGGCTCTGGAAT 3174
 QY 1001 ----- GluIleAlaGlnValIleAlaSerTyr 1009
 Db 3175 GCTGGGAAACAGGGAGTTTAGGAAAAAACCTGAAATTTGCCACGGTTATTCTCTCTAC 3234
 QY 1010 ThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuIleLeuArgLys 1029
 Db 3235 GCTGTACTGTGTCCTCCGAAACAACTCACTCCCTGGCTCTCTGGGCAGCTGATTCTGATCGGAA 3294

1030 LysAsnProGlyGlyTyrTrpGluGlyGluLeuGlnAlaArgGlyLysLysArgGlnIle 1049
 Db 3295 AAGAACCCAGTGGATGGTGGGAGGAGAACTCAAGCTCGAGGAAAAAGCGCCAGATA 3354
 QY 1050 GlyTyrPheProAlaAsnTyrValLysLeuLeuSerProGlyThrSerLysIleThrPro 1069
 Db 3355 GGGTGGTTTCCAGCAAAATATGTCAAACTTCTAAGCCCGGAAACAGCAAAATCACCCCA 3414
 QY 1070 ThrGluProProLysSerThrAlaLeuAlaValCysGlnValIleGlyMetTyrAsp 1089
 Db 3415 ACTGAGTACTCCCAAGACCCGAGTCGAGCCAGCTGTGCCAGGTGATCGGATGTACGAT 3474
 QY 1090 TyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLysGlyGlnIleIleAsnValLeu 1109
 Db 3475 TACACCGCCAGCAACGATGACGAATAGCTTTCAGCAAGAGCCAGATCATCAAGCTCTC 3534
 QY 1110 AsnLysGluAspProAspTyrTrpLysGlyGluValAsnGlyGlnValGlyLeuPhePro 1129
 Db 3535 AACAAAGGAGGACCGGACTGGTGGAAAGGAGATCAGTGGCAAGTTGGGCTCTTCCCA 3594
 QY 1130 SerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
 Db 3595 TCCAATATTGTAAAGCTGACCACAGATGAGACCCCGAGCCAG 3636

RESULT 6
 AAZ39025
 ID AAZ39025 standard; cDNA; 5144 BP.
 XX
 AC AAZ39025;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse EseiL coding sequence.
 XX
 KW Mouse; murine; Esei; Esei2; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 XX
 PN WO9955728-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-CA000375.
 XX
 PR 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Egan SE, Wang W, Sengar A;
 DR WPI; 2000-052802/04.
 DR P-PSDB; AAY57449.
 XX
 PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 PS Claim 6; Page 59-62; 99pp; English.
 CC
 CC The present invention specifically describes mammalian Esei and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp15 protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) or its (anti)agonists,
 CC mimetics, fragments and inactive mutants; (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Esei genes or antagonists of an Esei binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Esei is used

CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (1) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (1) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Bae mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Bae-Esp15 complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Bse11 coding sequence
 XX
 SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3, 56e-279 Length: 5144
 Score: 5437.50 Matches: 1059
 Percent Similarity: 90.28% Conservative: 37
 Best Local Similarity: 87.23% Mismatches: 45
 Query Match: 92.24% Indels: 73
 DB: 3 Gaps: 3

US-09-720-934-2 (1-1143) x AAZ39025 (1-5144)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
 DB 1 ATGGCTCAGTTTCCACACCTTTCGGTGGTAGCTCGATGCTGGGCCCATAACTGTGGAG 60
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 DB 61 GAAAGGGCCAGATGATCAGCAGCAGTCTCTTAGCTGAGCCGATAGCGGATTTTACT 120
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 DB 121 GGTGATCAAGCAGGAGAACTTTTTTCCAACTCTGGGTACCTCAGCCTGTCTTAGCACAA 180
 QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 DB 181 ATATGGGCGCTAGCGACATGAATACGATGGAGGATGGATCAAGTGAATTTTCCATA 240
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTrpGlnLeuProSerAlaLeuPro 100
 DB 241 GCCATGAAGCTTATCAATCTGAAGCTACAGAGATACAGCTCCCTCCACACTTCCCT 300
 QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 DB 301 GTCATGAAACAGCAACCACTGGCTATTTCAGTGCACCAAGCATTTGGTATAGAGGGATT 360
 QY 121 AlaSerMetProLeuThrAlaValSerValProThrAlaAlaValProLeuAlaAsn 140
 DB 361 GCTACATGCCACCACTCAGCTGTGTCTCTGTGCCAATGGGTCCCATTCAGTTGTT 420
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProLeuAlaAsn 160
 DB 421 GGAATGTCTCCACCCTTAGTATCTTCTGCTCCAGCAGCAGTGCCTCCCTGGCTAAC 480
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 DB 481 GGGGCTCTCCCGTCATACAGCTCTGCGTGGTGTGGCGATCTCGACCCACATGGCCA 540
 QY 181 LysSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 DB 541 AAGAGTTCTCTCTTACAGCAGATCTGGTCCAGGGTCACAAATTAACACTAAGTTACAGAG 600
 QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
 DB 601 GCACAATCATTCGATCTCCCGCGGCCCTCCAGCAGCAGATGGGCTGTGCTCAGTCA 660
 QY 221 SerArgLeuLysTyArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 DB 661 TCAAGGCTGAATATACAGGAGTATTTCACACGCCACGACAAACTATGATGGACACTTA 720
 QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260

DB 721 ACAGGTCCCAGGCAAGAACTATCTCATCAATCAAGTTTACCCAGGCTCAGCTGCT 780
 QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluPheIle 280
 DB 781 TCAATATGGAATCTTCTGACATTGATCAAGATGGAAATCTCACTGAGAAATTTATC 840
 QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
 DB 841 CTAGCTATGCACCTAATGATGTGCAATGTCTGGTCAGCCATCGCCCGCTCTGCT 900
 QY 301 ProGluTrpIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
 DB 901 CCAGAATACATCCCTCTCTTCCAGAGAGTTCTGCTCCGGCAGTGGGATGTCGTCATA 960
 QY 321 SerSerThrSerValAspGlnArgLeuProGluProValLeuGluAspGluGlnGln 340
 DB 961 AGCTCTTCTTCTGTGGATCAGAGGCTGCTGTGAGGAGCCCTGTCAGAGGATGAGCAG 1020
 QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
 DB 1021 CCA---GAGAGAACTGCTGTGACATTTGAAGATAGAGCGGAGRACTTCGAGCGA 1077
 QY 361 GlyAsnLeuGluLeuGlyLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
 DB 1078 GGCAGTGTGGAGTGGAGAGCGCCGCAAGCGCTCTTGAGCAGCAGCGCAAGAGCAG 1137
 QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluGlnGlu 400
 DB 1138 GAGCGTGTGGCTCAGCTGGAGCGCGCCGAGCAGGAGGAGGAGCGGAGCGCCAGGAG 1197
 QY 401 GlnGluArgLysArgGlnLeuGluLysGlnLeuLysGlnArgGluLeuGlu 420
 DB 1198 CAGGAGCCCAAGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCTGGAG 1257
 QY 421 ArgGlnArgGluGluArgLysGluIleGluArgGluAlaAlaLysArgGlu 440
 DB 1258 CGGACGAGAGGAGGAGGAGGAGGAGGAGATCGAGAGCGCGAGCGCCCAAAACGGGAA 1317
 QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGln 460
 DB 1318 CTGGAAGGCGAGCGCACTTGAATGGGAACCGAACCGAGACAGAACTCTCTGAATCAG 1377
 QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
 DB 1378 AGGAAACAAGCAGCAGCGAGCGCCCTGGTCTCTGAGGAGGAGGAGGAGCTCTGGAGTTT 1437
 QY 481 GluLeuGluAlaLeuAsnAspLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
 DB 1438 GAGTTAGAGCTCTGAATGACAAAAAGCATCAGCTAGAGAGGAAAACTTCAGGATATCAG 1497
 QY 501 CysArgLeuThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg 520
 DB 1498 TGTCCACTGGCAACCCAGAGCGAGAAATGAGACCGAACCAAGTCTAGAGAGCTAAGA 1557
 QY 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGlnGlnMetLeuGlyArg 540
 DB 1558 ATTGCTGAAATCACCCACTTACAGCAGCAGTGTGAGGAATCTCAGCAATGCTTGAAGA 1617
 QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
 DB 1618 CTATTTCCAGAGAAACAGATACTCAGTACCACCTGTTAAACAAAGTCCAGCAGAACAGTTTG 1677
 QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
 DB 1678 CATAGAGACTCTCTTCTTACCTCAAAAGAGCTTTGGAAGCAAGAGGAGTGGCCCGGAG 1737
 QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
 DB 1738 CAGCTCCGGAGCAGCTGGAGGAGTGGAGAGAGAGACAGGTCAGGTCAGAGGAT 1797
 QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620

1798 GATGTTTTCACACACAGCTGAGGACTGAGAGATACATAGCAAAACAGCAACTCCAG 1857
621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnArgLysIleIle 640
1858 AAGCAGAGTCCCTGAGGAGCGCGACTGAACGAAAGAGCAGGAGAGAGCGCTG 1917
641 GluLeuGlnLysGlnLysGluAlaGlnArgArgAlaGlnLysArgLysGlnTTP 660
1918 GAGTTAGAGAGCAAAAGGAAGACGCTCAGAGAGGAGTTGAGAAAGGCAAGCAATGG 1977
661 LeuGluHisValGlnGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGln 680
1978 CTGAGCATGTGAGAGAGAG---GAGCAGCCAGCGCCCGGAAACCCACGAGGAGGAC 2034
681 LysLeuLysArgGluGlnSerValLysLysLysLysLysLysLysLysLysLys 700
2035 AGACTGAAGAGGAGAGACAGTGTGAGAGAGAGGCGGAGAGAGAGCAAGCGGAA 2094
701 AlaGlnAspLysLeuGlnArgLeuPheHisGlnLysGlnLysGlnLysProAlaVal 720
2095 ATGCAACAGACAGAGTGGCTTTTCCATCCGATCAGGAGCCAGTAAAGCTGGCCACC 2154
721 GlnAlaProTyrSerThrAlaGluLysGlnLysGlnLysGlnLysGlnLysGln 740
2155 CAGGACCCCTGCTCTACCAAGAGAGAGGCGCGCTTACCATTTCTGCACAGGAGAGTGA 2214
741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
2215 AAAGTGTATATACCGAGCGCTGATCCCTTTGAATCCAGAGTACGATGAGATCACC 2274
761 IleGlnProGlyAspIleValMetValAspLysSerGlnThrGlyGluProGlyTyrPhe 780
2275 ATCCAGCCAGGAGATATAGTCATGCTGATGATGAAGCCAGACTGAGAGCCAGGATGCTT 2334
781 GlyGlyGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 800
2335 GGAGGAGAGTGAAGGAGAGAGCGGATGTTCTCTGCAACTATGAGAGAGAGATTCGA 2394
801 GluAsnGlnValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro 820
2395 GAAATGAGGTTCCTCCAGCCAAACAGTGCAGTCCGATCTGACCTCGCCCTCGCCCC 2454
821 LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerGluProSerThr 840
2455 AAATGCTCTGCTGAGAGCCCTCTCTTTCGAGAGAGCTCTCTGAGCCCTCCACA 2514
841 ThrProAsnAsnTyrAlaAspPheSerSerThrTyrProThrSerThrAsnGluLysPro 860
2515 ACCCCCAACAACCTGGGAGAGCTTCAGTTCCAGTGGCCAGCAGCTCAACGAGAGCCA 2574
861 GluThrAspAsnTyrAspAlaTyrAlaGlnProSerLeuThrValProSerAlaGly 880
2575 GAAACGGCAACTGGGATACGTGGCGGCTCAGCCCTCTCTGACCGTACCTAGTGTCTGGC 2634
881 GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro 900
2635 CAGTTACGGCAGAGATCAGCTTTTACCCAGCCACAGCCACTGGCTCTCTCCCATCTCC 2694
901 ValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaGlnAlaLeuTyrProTyrArg 920
2695 GTCCTGGCCAGGGTGAAGAGGTGAAGGCTTACAAGGCGCAAGCCCTGTATCCCTGGAGA 2754
921 AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGln 940
2755 GCCAAAAAGACAACCTTAATTTTAAACAAAGTACGTCATCCCTCTCTGGAACAG 2814
941 GlnAspMetTyrTyrPheGlyGluValGlnGlnLysGlnLysGlyTyrPheProLysSerTyr 960
2815 CAAGACATGTGTGTGGAGAGTTCAAGTCAAGAGGGTTGGTTCCCAAGTCTTAC 2874
961 ValLysLeuLysSerGlyProLeuArgLysSerThrSerMetAspSerGlySerSerGlu 980
2875 GTGAACTCATTTGAGGCGCGGTAGGAAATCCAAAGATCATGATCGCCCTACTGAA 2934

981 SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu 1000
2935 AGTCTCTGCTAGTCTAAAGAGAGTGGCTTCCCGCGCGCCAAAGCCAGCATTCGCGAGAA 2994
1000 ----- 1000
2995 GAGTTTATTTGCCATGTACACATACGAGAGTCTCTGACAAAGGAGATTAACTTTTCAGCAA 3054
1000 ----- 1000
3055 GGGGATGTGATTGTGTTTACCAAGAAAGATGGTGACTGTGTGACGGGAACGGTGGCGCAC 3114
1000 ----- 1000
3115 AAGTCCGGAGTCTTCCCTTCTTAATATGTGAGGCTTTAAAGATTTCAGAGGCTCTGGAACT 3174
1001 ----- GluIleAlaGlnValIleAlaSerTyr 1009
3175 GCTGGGAAAACAGGGAGTTTAGGAAAAAACCTGAAATTCGCCAGGTTATTGCTTCTCTAC 3234
1010 ThrAlaThrGlyProGlnLeuThrLeuAlaProGlyGlnLeuIleLeuIleArgLys 1029
3235 GCTGCTACTGTCCTCCGAAACAACTCACCTGCTCTCTGGGAGCTGATTCTGATCCGAAA 3294
1030 LysAsnProGlyGlyTyrTyrGluGlnAlaArgGlyLysLysLysArgGlnIle 1049
3295 AAGAACCCAGTGTGATGGTGGAGAGAACTGCAAGCTCGAGGAAAAAGCGCCAGATA 3354
1050 GlyTyrPheProAlaAsnTyrValLysLeuLeuSerProGlyThrSerLysValIleThrPro 1069
3355 GGGTGGTTTCCAGCAATATATGTCAAACCTTCTAAGCCCCGGAACAAGCAAAATCACCCCA 3414
1070 ThrGluProProLysSerThrAlaLeuAlaValCysGlnValIleGlyMetTyrAsp 1089
3415 ACTGAGTACCCAGACCGCAGTGCAGCCAGCAGTGTGCCAGGTGATCGGATGTACGAT 3474
1090 TyrThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGlnIleIleAsnValLeu 1109
3475 TACACCGCCAGAACAGATGACAACTAGCCTTCAGCAAGGCGCAGATCATCAACGTCTCTC 3534
1110 AsnLysGluAspProAspTyrTyrLysGlyGlnValAsnGlyGlnValGlyLeuPhePro 1129
3535 AACAGGAGGAGCCCGAGCTGGTGGAAAGGAGAGTCAAGTGGGCAAGTTGGGCTCTTCCCA 3594
1130 SerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
3595 TCCAAATTATGTAAAGCTGACCAAGACATGACCCCGCCAG 3636

RESULT 7

AAZ39024
ID AAZ39024 standard; cDNA; 5738 BP.

AC AAZ39024;

DT 28-FEB-2000 (first entry)

DE Mouse EseiL cDNA sequence.

XX Mouse; murine; Esei; Esei2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.

OS Mus sp.

XX WO9955728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 98CA-02230201.

PR 05-FEB-1999; 99US-0118739P.
 XX (HSCR-) HSC RES & DEV LP.
 XX Egan SE, Wang W, Sengar A;
 XX WPI; 2000-052802/04.
 DR P-PSDB; AAY57449.
 XX
 XX New nucleic acid encoding Eesl and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 XX Claim 6; Page 56-59; 99pp; English.
 XX
 CC The present invention specifically describes mammalian Eesl and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp15 protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Eesl is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Eesl cDNA sequence
 XX

Alignment Scores:
 Pred. No.: 4, 02e-279 Length: 5738
 Score: 5437.50 Matches: 1059
 Percent Similarity: 90.28% Conservative: 37
 Best Local Similarity: 87.23% Mismatches: 45
 Query Match: 92.24% Indels: 73
 DB: 3 Gaps: 3

US-09-720-934-2 (1-1143) x AAZ39024 (1-5738)

QY 1 MetAlaGlnPheProThrPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
 DB 260 ATGGCTCAGTTTCCACACCTTTCGGTGTAGCTGGATGTCGGGCCATACTGTGGAG 319
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 DB 320 GAAAGGGCCCAAGCATCACCAGCAGTTCCTTAGCTCGAAGCCGATAGCGGGATTATTACT 379
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 DB 380 GGTGATCAAGCAGGAGGAACCTTTTTCCTCAATCTGGGTATCCTCAGCCTGTCTTAGCACA 439
 QY 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 DB 440 ATATGGCGCTAGCGGACATGATACGATGAGGATGATCAAGTGGAAATTTCCATA 499
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 DB 500 GCCATGAAGCTTATCAAACTGAAGCTACAGGATATCAGTCCCTCCCATCTCCCTCCCT 559
 QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 DB 560 GTCATGAACACAGCAGCAGTGGCTATTTCAGTGCACACGACATTTGGTATAGGAGGATT 619
 QY 121 AlaSerMetProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140

DB 620 GCTAGCATGCCACCACCTCACAGCTGTGTCTCTGTGCCAATGGCTCCATCCAGTTGT 679
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 DB 680 GGAATGTCTCCACCTTAGTATCTTGTCTCCCTCCAGCAGCAGTCCCTCCCTCGCTAAC 739
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 DB 740 GGCGCTCTCCCTCCGTCATACAGCCTCTGCTGCGTGTGGCATCTCTGACGCCACATGGCCA 799
 QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 DB 800 AAGAGTTCTCTTCCTTCAGCAGATCTGGTCCAGGGTCACAATTTAAACACTAAGTTACAGAAG 859
 QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
 DB 860 GCACAATCATTCGATGCGCAGCGCCCTCCAGCAGCAGCAATGGGCTGTGCTCAGTCA 919
 QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 DB 920 TCAAGGCTGAATACAGCAGTATTTCACAGCCACACCAAACTATAGTGGACACTTA 979
 QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
 DB 980 ACAGGTCCTCCAGGCAAGAACTATTCTCATGCAATCAAGTTTACCCAGGCTCAGCTGGCT 1039
 QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
 DB 1040 TCAATATGGAATCTTCTGACATTGATCAAGATGGAAACTCACTGCAGAAGAAATTTATC 1099
 QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
 DB 1100 CTAGCTATGACCTAATTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
 QY 301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
 DB 1160 CCAGAATACATCCCTCTCTTCCAGAGAGTTTCGCTCCGGCAGTGGGATGTCGTCATA 1219
 QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
 DB 1220 AGCTCTTCTTCTGTGATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1279
 QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
 DB 1280 CCA--GAGAAGAACTGCCTGTGACATTGAAGATAGAGCGGAGAACTTCGAGCGGA 1336
 QY 361 GlyAsnLeuLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
 DB 1337 GGCAGTGTGGAGCTGGAGAGCGCCGCAAGCGCTCTTGAGCAGCAGCGCAAGAGCAG 1396
 QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnArgLysGluArgGluArgGlnGlu 400
 DB 1397 GAGCGTTTGGCTCAGCTGGAGCGCGCCGAGCAGGAGAGAAAGCGGAGCGCCAGGAG 1456
 QY 401 GlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
 DB 1457 CAGGAGCCCAAGCGCAGCTGGAGCTGGAGAGCAGCTGGAGAGCAGCGGAGCGGAGCTGGAG 1516
 QY 421 ArgGlnArgGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu 440
 DB 1517 CGGACGAGAGAGGAGAGGAGGAGAGAGATCGAGAGCGCGAGCGCGCAAAACGGGAA 1576
 QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnLeuLeuAsnGln 460
 DB 1577 CTGGAAAGCAGCAGCAACTTGAATGGGAACCGAACCGGAGACAGAACTCTCGATCAG 1636
 QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
 DB 1637 AGGAACAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1696
 QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500

Db	1697	GAGTTAGAAAGCTCTGAATGACAAAAGACATCAGCTAGAGGAAAACCTTCAGGATATCAGG	1756	861	GluThrAspAsnTrpAspAlaTrpAlaGlnProSerLeuThrValProSerAlaGly	880
QY	501	CysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg	520	2834	GAACCGACAACTGGGATACGTGGGGGGCTCAGCCTTCTGTGACCGTACCTAGTGTGGC	2893
Db	1757	TGTCGACTGGCAACCCAGGCAAGAAATGAGAGCACGAACAAGTCTAGAGAGCTAAGA	1816	881	GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerProSerPro	900
QY	521	IleAlaGluIleThrHisLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	540	2894	CAGTTACGGCAGAGATCAGCCTTTACCCAGCCAGCAGCCACTGGCTCTCCCATCTCCC	2953
Db	1817	ATTGCTGAATCACCACCTTACAGCAGCAGTTCAGGAATCTCAGAAATGCTTGGAGA	1876	901	ValLeuGlyGlnGlyGlnGlyValGluGlyLeuGlnAlaGlnAlaLeuThrProTrpArg	920
QY	541	LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu	560	2954	GTCTTGGCCAGGGTGAAGAAGGTGGAAGGGCTACAAAGCGCAAGCCCTGTATCCCTGAGA	3013
Db	1877	CTTATTCCAGAGAAACAGATCTCAGTGACCACTTAAACAGTCCAGCAGAAACAGTTTG	1936	921	AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGlnGln	940
QY	561	HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln	580	3014	GCCAAAGAACACCAACCTTAAATTTTAAACAAAGTGACGTATCATCCGTTCTGGAACAG	3073
Db	1937	CATAGAGACTCGCTCTTACCTCAAAGAGCCTTGGAGCAAAAGGAGCTGGCCCGCAG	1996	941	GlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrpPheProLysSerTyr	960
QY	581	HisLeuArgAspGlnLeuAspGluValGluLysGlnThrArgSerLysLeuGlnGluIle	600	3074	CAAGACATGT	3133
Db	1997	CAGCTCCGGGAGCAGCTGGACGAGTGGAGAGAGAGACCAAGCTCAAAAGCTGACGAGATT	2056	961	ValLysLeuIleSerGlyProIleArgLysSerThrSerMetAspSerGlySerSerGlu	980
QY	601	AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln	620	3134	GTGAACTCATTTTCAGGCCCCGTAAAGGAATCCACAAGCATCGATACTGGCCCTACTGAA	3193
Db	2057	GATGTTTTCAACACCACTGAGGAAGTCTGAGAGATACATATCAAAAGCAACTCCAG	2116	981	SerProAlaSerLeuLysArgValAlaSerProAlaLysProValValSerGlyGlu	1000
QY	621	LysGlnLysSerMetGluAlaGluArgLysGlnLysGlnGlnGlnGlnGlnGlnGln	640	3194	AGTCTCTAGTCTTAAAGAGAGTGGCTTCCCGCCGCGCAAGCCAGCATTCCTCCGGAGAA	3253
Db	2117	AAGCAGAGGTCCTGGAGCAGCGACTGAAGCAGAAAGAGCAGGAGGAGGAGGAGCTG	2176	1000	-----	1000
QY	641	GluLeuGluLysGlnLysGluAlaGlnArgAlaGlnGluArgAspLysGlnTrp	660	3254	GAGTTTATTGCAATGTACATACGAGAGTTCTGAGCAAGAGAGATTAACTTTTCAGCAA	3313
Db	2177	GAGTTAGAGAAAGAAAGAGCGCTCAGAGACGAGTTTCAGGAAAGGGAAGAAGATGG	2236	1000	-----	1000
QY	661	LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu	680	3314	GGGATGTGATGT	3373
Db	2237	CTGAGCATGTGACAGAGAG---GAGCAGCAGCCCGCCCGGAAACCCACGAGGAGAC	2293	1000	-----	1000
QY	681	LysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLysGlyLysGlnGlu	700	3374	AAGTCCGGAGTCTTCCCTTCTAACTATGTGAGGCTTAAAGATTTCAGAGGGTCTTGGAACT	3433
Db	2294	AGACTGAAGAGGAAGACAGTGTACAGGAAGAGGAGGCGGAAGAGAGAGCCAAAGCCGAA	2353	1001	-----	1001
QY	701	AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal	720	3434	GCTGGGAAAAACAGGAGGTTTAGGAAAAAAACCTGAAATTTGCCAGGTTATTGCTTCTTAC	3493
Db	2354	ATGCAAGACAAACAGAGTCGGCTTTTCCATCCGATCAGGACCCAGCTAAGTGGCCACC	2413	1010	ThrAlaThrGlyProGluGluLeuThrLeuAlaProGlyGlnLeuLeuLeuLeuLeuLys	1029
QY	721	GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal	740	3494	GCTGTCTGTGTCGGAACAACTCACCTGGCTCTGGGCGAGTGTGATTCGATCCGGA	3553
Db	2414	CAGGCACCTGCTTACCACAGAGAAAGGCCGCTTACCATTCTGCACAGGAGAGTGA	2473	1030	LysAsnProGlyGlyTrpTrpGluGluLeuGlnAlaArgGlyLysLysArgGlnIle	1049
QY	741	LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr	760	3554	AAGAACCCAGGTGGATGTGGGAGGAGAACTGCAAGCTCGAGGGAAGAAAGCCAGATA	3613
Db	2474	AAAGTGGTATATTACCGAGCGCTGTACCCCTTTGAATCCAGAAAGTCAAGATCAGATCAC	2533	1050	GlyTrpPheProAlaAsnTyrValLysLeuLeuSerProGlyThrSerLysIleThrPro	1069
QY	761	IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeu	780	3614	GGGTGGTTCAGCAAAATTTATGTCAAACTTTAAGCCCGGGAACAGCAAAATCACCCCA	3673
Db	2534	ATCCAGCCAGGAGATATAGTCTATGGTGGATGAAGCCAGACTGGAGAGCCAGATGGCTT	2593	1070	ThrGluProProLysSerThrAlaLeuAlaValCysGlnValIleGlyMetTyrAsp	1089
QY	781	GlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIlePro	800	3674	ACTGAGCTACCCCAAGCCAGTCGAGCCAGCGAGTGTGCGAGGTATCGGATGTACGAT	3733
Db	2594	GGAGGAGACTGAAGGAGAGCGGATGTTTCCCTGCAAACTATCGAAGAAAGATTCCA	2653	1090	TyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLysGlyGlnIleIleAsnValLeu	1109
QY	801	GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro	820	3734	TACACCGCCAGAACCATGACAACTAGCCTTCAGCAAAAGGCCAGATCATCAACGTCCTC	3793
Db	2654	GAAATATGAGTTCCTCCACTCCAGCCAAACAGTACCGATCTGACATCTGCCCTGCCCC	2713	1110	AsnLysGluAspProAspTrpTrpLysGlyGluValAsnGlyGlnValGlyLeuPhePro	1129
QY	821	LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr	840	3794	AACAAGAGGAGACCCGAGCTGGTGGAAAGAGAGAGTCTAGTGGGCAAGTGGGCTCTTCCCA	3853
Db	2714	AAACTGGCTCTGGGTGAGACCCCTGCTCTTCCAGTAGTACCTTCTTCCAGCCCTCCACA	2773	1130	SerAsnTyrValLysLeuThrThrAspMetAspProSerGln	1143
QY	841	ThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysPro	860	3854	TCCAAATATGTAAGCTGACCAACAGACATGGACCCCGAGCAG	3895
Db	2774	ACCCCAACAACTGGGAGACTTCAGTTCACGCTGGCCCGCAGCAGCTCAAAACGAGAAAGCA	2833			

RESULT 8

AAZ39008
ID AAZ39008 standard; cDNA; 5082 BP.

XX AAZ39008;

XX 28-FEB-2000 (first entry)

XX Mouse Esei full length cDNA sequence.

XX Mouse; murine; Esei; Esei2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.

XX Mus sp.

XX WO995728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 98CA-02230201.

PR 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

XX Egan SE, Wang W, Sengar A;

XX WPI; 2000-052802/04.

DR P-PSDB; AAY57444.

XX New nucleic acid encoding Esei and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.

XX Claim 6; Page 38-40; 99pp; English.

XX The present sequence encodes mouse Esei. The present invention
CC specifically describes mammalian Esei and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esep5 protein) vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Esei is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esep5 complex, then binding dynamitin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX

SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,22e-277	Length:	5082
Score:	5408.50	Matches:	1058
Percent Similarity:	90.20%	Conservative:	37
Best Local Similarity:	97.15%	Mismatches:	46
Query Match:	81.75%	Indels:	74
DB:	3	Gaps:	3

US-09-720-934-2 (1-1143) x AAZ39008 (1-5082)

QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20

Db	259	ATGGCTCAGTTTCCACACACCTTTTCGGTGGTAGCTGCTGGGCATTAACCTGGAG	318
QY	21	GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr	40
Db	319	GAAAGGGCCAGCATGACCGAGTTCCTTAGCCCTGAAGCCGATAGCGGGATTATTACT	378
QY	41	GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln	60
Db	379	GGTGATCAAGCGAGAACTTTTTCATCTCGGTACCTCAGCCCTGCTTAGACAA	438
QY	61	IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle	80
Db	439	ATATGGCGCTAGCGGACATGAATAACGATGAAGGATGATCAAGTGAATTTCCATA	498
QY	81	AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro	100
Db	499	GCCATGAAGCTTATCAAACTGAAGCTACAAGGATATCAGCTCCCTCCACACTTCC	558
QY	101	ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle	120
Db	559	GTCAATGAACAGCAACCCAGTGGCTATTTCAGTGCACCGACATTTGGTATAGGAGG	618
QY	121	AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal	140
Db	619	GCTAGCATGCGCAGCACTCAGAGCTGTGTCTCTGTGCCAATGGGCTCCATTCCAGT	678
QY	141	GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn	160
Db	679	GGATGTCCTCACCTTAGTATCTTCTGCTCCCTCCAGCAGCAGTGCCTCCCTGGCT	738
QY	161	GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro	180
Db	739	GGGGCTCTCCCTCATACAGCTCTGCTGCGTGGCATCTCTGCGCAGCCATGCGCCA	798
QY	181	LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys	200
Db	799	AAGAGTTCTTCCTTCAGCAGATCTGGTCCAGGGTCACAATTAACACATAAGATTAC	858
QY	201	AlaGlnSerPheAspValAlaSerValProProValAlaGlnTrpAlaValProGlnSer	220
Db	859	GCAATCATTCAGTTCGCCAGCGCCCTCCAGCAGCAGAGATGGGCTGTGCTCAGTCA	918
QY	221	SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu	240
Db	919	TCAAGGCTGAAATACAGGCGATTATTCACAGCCAGCAGCAAACTATGATGGACACTTA	978
QY	241	ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla	260
Db	979	ACAGGTCCCGCAGCAAGAACTATTCATGCAATCAAGTTTACCCAGGCTCAGTGGCT	1038
QY	261	SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluPheIle	280
Db	1039	TCATATGAAATCTTCTGACATTGATCAAGATGAAGAACTACTCGAGAAAGATTATC	1098
QY	281	LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro	300
Db	1099	CTAGCTATGACCTAATTGATGTGGTTCATGCTGGTCAGCCATGCGCGCCCTCTGCCT	1158
QY	301	ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle	320
Db	1159	CCAGAATACATCCCTCTCTTCAGAGAGTTCGCTCCGGCAGTGGGATGTCGGTCATA	1218
QY	321	SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln	340
Db	1219	AGCTCTTCTTGTGGATCAGAGGGTGCCTGAGGAGCCGCTCGTAGAGATGAGCAGCAG	1278
QY	341	GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg	360
Db	1279	CCA--GAGAAGAAACTGCTGACATTTGAAGATGAAGCGGGAGAACTTCGAGCGCA	1335
QY	361	GlyAsnLeuGluLeuLysArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln	380

1336 GGAGTGTGAGCTGAGAAAGCCGCCAAGCGCTCTTGAGCAGCAGCGCAAGAGCAG 1395
 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluGlnGlu 400
 1396 GAGCGTGGCTCAGCTGGAGCGCGCGAGCAGAGAGAAAGAGCGGAGCGCCAGGAG 1455
 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnLeuGlu 420
 1456 CAGGAGCCCAAGCGCAGCTGGAGCTGGAGAAGCAGCTGGAGAAGCAGCGGAGCTGGAG 1515
 421 ArgGlnArgGluGluArgArgLysGluLeuGluArgGluAlaAlaLysArgGlu 440
 1516 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1575
 441 LeuGluArgGlnArgGlnLeuGluLysGlnLeuGluArgGlnGlnLeuLeuAsnGln 460
 1576 CTGGAAAGCGAGCGACAACTTGAATGGAAACCGGAAACCGGACAGGAACTCTGAATCAG 1635
 461 ArgAsnLysGlnGlnGluAspIleValLeuLysAlaLysLysLysLysLysLysLys 480
 1636 AGGAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1695
 481 GluLeuGluAlaLeuAsnAspLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
 1696 GAGTTAGAGCTCTGAATGACAAAGATCAGCTAGAGGAAACATTCAGGATATCAGG 1755
 501 CysArgLeuThrGlnArgGlnGlnGluLeuSerThrAsnLysSerArgGluLeuArg 520
 1756 TGTCGACTGGCAACCCAGAGCAAGAAATTGAGAGCAGCAACAAAGTCTAGAGAGCTAAGA 1815
 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArg 540
 1816 ATTGCTGAATTCACCCACTTACAGCAGCAGGTTGAGGAAATCTCAGCAAAATGCTTGGAGA 1875
 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
 1876 CTTATTCCAGAAACACAGATACTCAGTCACCCAGTTTAAACAAAGTCCAGCAGCAACAGTTTG 1935
 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
 1936 CATAGAGACTCGCTCTTACCTCAAAAGAGCCTTGGAAAGCAAGAGCCTGGCCCGGAG 1995
 591 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
 1996 CAGCTCCCGGAGCAGCTGGACAGGTGGAGAGAGAGACAGCTCAAGCTCAGCAGAGATT 2055
 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
 2056 GATGTTTCAACCAACAGCTGAAGCACTGAGAGAGATACATAGCAACAGCAACTCCAG 2115
 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGlnGluArgLysIleIle 640
 2116 AAGCAGAGGTCCTGGAGCAGCGGACTGAAGCAGAAAGAGCAGGAGGAGGAGGAGCCTG 2175
 641 GluLeuGluLysGlnLysGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrp 660
 2176 GAGTTAGAGCAACAAAGAGAGACCTCAGAGCAGCTCAGAAAGGAGCAAGCAATGG 2235
 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
 2236 CTGGAGCATGTGCAGCAGGAG--GAGCAGCCACGCCCGCCGAAACCCACGAGGAGGAG 2292
 681 LysLeuLysArgGluGluSerValLysLysLysLysLysLysLysLysLysLysLys 700
 2293 AGACTCAAGAGGAGGAGAGCAGTGTCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2352
 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGlnProAlaLysProAlaVal 720
 2353 ATGCAAGACAGCAGAGTGGCTTCCATCCGATCAGCAGCAGCTAAGCTGGCCACC 2412
 721 GlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
 2413 CAGGCAACCTGGTCTACACAGAGAAAGCCCGCTTACCATTCTGCACAGGAGAGTGTA 2472

741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
 2473 AAGTGGTATATTACCGAGCGCTGTACCCCTTGAATCCAGAAGTCACGATGAGATCACC 2532
 761 IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeu 780
 2533 ATCCAGCCAGGAGATATAGTCATGCTGATGAAGCCAGACTGGAGAGCCAGGATGGCTT 2592
 781 GlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIlePro 800
 2593 GAGAGGAGCTGAAGAGGAGAGCGGATGGTTCCTGCAAACTATGCAAGAAAGATTCCA 2652
 801 GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro 820
 2653 GAAATAGCTTCCCATCTCCAGCCAAACAGTCAGCATCTGACCCCTGCCCC 2712
 821 LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr 840
 2713 AAACCTGCTCTGCTGAGACCCCTGCTCTTGGCAGTGACCTCTTCTGAGCCCTCCACA 2772
 841 ThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysPro 860
 2773 ACCCCCAACACTGGCAGACTTCAGTTCACGTGGCCAGCAGCTCAACAGAGAGCCA 2832
 861 GluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThrValProSerAlaGly 880
 2833 GAAACGACCAACTGGAT-ACGTGGCGGCTCAGCCTTCTCTGACCGTACCTAGTGGTGC 2891
 881 GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro 900
 2892 CAGTTACGCGCAGAGATCAGCCTTTACCCAGCCAGCCACTGGCTCTCTCCCATCTCCC 2951
 901 ValLeuGlyGlnGlyLysValGluLysValGluLysAlaGlnAlaLeuTyrProTrpArg 920
 2952 GTCCTGGCCAGGTGAAAGGTGGAGGCTTACAGCGCAAGCCCTGATCCCTGGAGA 3011
 921 AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGln 940
 3012 GCCAAAAAGACAACCACTTAAATTTTAAACAAAGTACGCTCATCCGCTTCTGGAACAG 3071
 941 GlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrpPheProLysSerTyr 960
 3072 CAAGACATGTGTGGTGGTGGAGAGTTCAGGTCAAGAGGTTGGTTCCTCCCAAGTCTTAC 3131
 961 ValLysLeuIleSerGlyProIleArgLysSerThrSerMetAspSerGlySerSerGlu 980
 3132 GTGAAACTCATTTGAGGCGCGTAAGGAAATCCACAGCATCGATCTGGCCCTACTGAA 3191
 981 SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu 1000
 3192 AGTCTGCTAGTCTAAAGAGAGTGGCTTCCCGCGCGCAAGCCAGCCATTCCTCCGAGAA 3251
 1000 ----- 1000
 3252 GAGTTATTGCGCATGTACATACAGAGATTCTGAGCAAGGAGATTAACTTTCAGCAA 3311
 1000 ----- 1000
 3312 GGGGATGTGATTGTGTTTACCAGAAAGATGGTGTGAGTGGTGGACGGGACGGTGGCGAC 3371
 1000 ----- 1000
 3372 AAGTCCGGAGTCTTCCCTTTAACTATGTAGGCTTAAAGATTACAGGGCTCTGGAAT 3431
 1001 ----- GluIleAlaGlnValIleAlaSerTyr 1009
 3432 GCTGGGAAACAGGAGGTTTAGGAAAAAACCTGAAATTTGCCAGGTTATTGCTTCTTAC 3491
 1010 ThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuIleLeuLys 1029
 3492 GCTGCTACTGGTCCCGAAACAACTCACCTGGCTCTCTGGGAGCTGATTCTGATCCGAAA 3551

PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246530P.
PR 08-NOV-2000; 2000US-0246531P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Barash SC, Ruben SM;
XX PT P-PSDB; AAU87168.
XX DR WPI; 2001-581633/65.
XX DR P-PSDB; AAU87168.
XX PT New isolated nucleic acid encoding a protein for diagnosing, preventing, or
XX PT treating or ameliorating medical conditions and used as food additives or
XX PT preservatives.
XX PT
XX PT
XX PS Claim 1: SEQ ID NO 88; 837pp; English.
XX CC The invention describes an isolated nucleic acid molecule (I) encoding a
XX CC novel central nervous system protein. (I) and polypeptides (III) encoded
XX CC by (I), are used to treat a medical conditions and in diagnosis of a
XX CC pathological condition. Disorders which are diagnosed or treated include
XX CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX CC disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
XX CC angiotensin, nervous system disorders e.g. Alzheimer's disease and
XX CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
XX CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
XX CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
XX CC adenocarcinomas and irritable bowel syndrome, reproductive system
XX CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
XX CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
XX CC leukaemia, disorders involving neovascularisation e.g. malignancies,
XX CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.

CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Alignment Scores: 9.56e-236 Length: 3319
Pred. No.: 4615.00 Matches: 922
Score: 91.81% Conservative: 9
Percent Similarity: 90.93% Mismatches: 27
Best Local Similarity: 78.29% Indels: 56
Query Match: 4 Gaps: 4
DB: 4
US-09-720-934-2 (1-1143) x ABK43498 (1-3319)
QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 159 ATGGCTCAGTTTCCACACACCTTTTGGTGGCAGCTGGATATCTGGGCCATAACTGTAGAG 218
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 219 GAAAGAGCGAAGCATGATCAGCGATTCCATAGTTTAAAGCCATATCTGGATTCTACT 278
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 279 GGTGATCAAGCTAGAAAACCTTTTTTCAATCTGGTTTACCTCAACCTGTTTATGACAG 338
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyVargMetAspGlnValGluPheSerIle 80
Db 339 ATATGGGCACCTAGTGCATGAATGATGGAAGAATGGATCAAGTGGAGTTTCCATA 398
QY 81 AlaMetLysLeuIleLysLeuLysGlnGlyVrGlnLeuProSerAlaLeuProPro 100
Db 399 GCTATGAAACTTATCAAACTGAAGCTACAGGATATCAGCTACCTCTGCATCTCCCT 458
QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db 459 GTCATGAAACAGCAACACCTGTTGCTATTTCTAGCGCACCAGCATTTGGTATGGAGGTATC 518
QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db 519 GCCAGCATGCCACCGCTTACAGCTGTTGCTCAGTGCACCAATGGATTCATTCAGTTGTT 578
QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db 579 GGAATGCTCCACACCTAGTATCTCTGTTCCACAGCAGCTGTGCCCCCTGGCTAAC 638
QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db 639 GGGGCTCCCCCTGTTTATACACCTCTGCTGATTTGCTATCTGTCAGCCACATGCGCA 698
QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db 699 AAGAGTTCTCTCTTAGTAGATCTGGTCCAGGGTCACAACTAAACACTAAATTCACAAAG 758
QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
Db 759 GCACAGTCATTGATGTGGCCAGTCCACCAGTGGCAGAGTGGGCTGTTCTCTCAGTCA 818
QY 221 SerArgLeuLysTyArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db 819 TCAAGACTGAAATACAGGCAATTTATTCATAGTAGTCATGACAAAACACTATGATGAGTGACACTTA 878
QY 241 Thr-----GlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGln 256
Db 879 ACAGGTCTCTGTTAGTCCCCAGCAAGAACTATTTCTTATGACAGTCAAGTTTACCACAG 938
QY 257 AlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAla 276
Db 939 GCTCAGCTGGCTTCMATATGGAATCTTCTGACATTTGATCAAGATGGAACACTTACAGCA 998

QY 277 GluGluPheIleLeuAlaMetHisLeuIleAspValalaMetSerGlyGlnProLeuPro 296
DB 999 GAGCAATTTATCTGCGCAATGCACTCAITGATGTAGTCTGCGCAACCACTGCCA 1058
QY 297 ProValLeuProGluTyrIleProProSerPheArgArgValArgSerGlySerGly 316
DB 1059 CCTGTCTGCTCCAGAAATACATTCACCTTCCTTTAGAAAGATTGATCTGCGAGTGT 1118
QY 317 IleSerValIleSerThrSerValAspGlnArgLeuProGluGluProValLeuGlu 336
DB 1119 ATATCTGTATAGTCAACATCTGTAGATCAGAGGCTACAGAGNACCAGTTTAGAA 1178
QY 337 AspGluGlnGlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGlu 356
DB 1179 GATGACAAACAATTTAGAAAAGAAATACCTGTAACTGTAACCTTTGAAGTAAGACGGAG 1238
QY 357 AsnPheGluArgGlyAsnLeuGluLeuGluLysArgGlnAlaLeuGluGlnGln 376
DB 1239 AACTTTGAACGTGGCAACCTGGAACTGGAGAAACGAAGGCAAGCTCTCTGGAACAGAG 1298
QY 377 ArgLysGluGlnGlnArgLeuAlaGlnLeuGluArgAlaGluGlnGlnArgLysGluArg 396
DB 1299 CGCAAGACAGAGAGCGCTGCGCCAGCTGAGCGCGGCGAGCAGAGAGAGAGCGT 1358
QY 397 GluArgGlnGlnGlnArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGln 416
DB 1359 GAGCGCCAGAGCAGAGCGCAAGAGCAACCTGGAACCTGGAAGCAACTGGAAGAGCAG 1418
QY 417 ArgGluLeuGluArgGlnArgGluGluArgLysGluIleGluArgGluAla 436
DB 1419 CCGGAGCTAGAACGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1478
QY 437 AlaLysArgGluLeuGlnArgGlnLeuGluTyrGluArgAsnArgArgGlnGlu 456
DB 1479 GCBAACCGGAACTTGAAGGCCAACGCAACTTGAGTGGAAACGGAATCGAAGGCAAGAA 1538
QY 457 LeuLeuAsnGlnArgAsnLysGluGlnGluAppIleValLeuLysLysLysLys 476
DB 1539 CTACTAAATCAAGAAACAAAGAACAAAGAGGACATAGTTGTACTGAAAGCAAAAGAAAG 1598
QY 477 ThrLeuGluPheGluLeuGluAlaLeuAsnAspLysHisGlnLeuGluLysLeu 496
DB 1599 ACTTTGSAATTTGAATTAGAGCTCTTAATGATAAAAGCATCACTAGAGGGGAACTT 1658
QY 497 GlnAspIleArgCysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSer 516
DB 1659 CAAGATATCAGATGTCGATTGACCCCAAGAGCAAGAAATTCAGAGCACAAACAAATCT 1718
QY 517 ArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnSerGlnGln 536
DB 1719 AGAGAGTTGAGAAATTCGCGAAATCACCCATCTACAGCAACAATTCAGAGAAATCTCAGCAA 1778
QY 537 MetLeuGlyArgLeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGln 556
DB 1779 ATGCTTGGAACTTATTCAGAAAACAGATCTCAATGACCAATTAAGAAACAGATTTCAG 1838
QY 557 GlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGlu 576
DB 1839 CAGAACAGTTTGCACAGAGATTCTGTTGTTACACTTAAAGAGCCTTAGAGCAAAAGAA 1898
QY 577 LeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLys 596
DB 1899 CTAGCTCGCAGCAGCCTACAGACCAACTGGATGAGTGAAGTGAAGAAATAACAAATAAG 1958
QY 597 LeuGlnGluIleAspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisLysLys 616
DB 1959 CTACAGGAGATTGATATTTTCAATATACAGTGAAGAACTAAGAGAAATAACAAATAAG 2018
QY 617 GlnGlnLeuGlnLysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGlu 636
DB 2019 CAACAACCTCCAGAGCAAAAGTCCATGGAGGCTGACAGCTGAAACAGAAAGAACAGAA 2078

QY 637 ArgLysIleIleGluLeuGluLysGlnLysGluAlaGlnArgArgAlaGlnGluArg 656
DB 2079 CGAAAGATCATAGAAATTAGAAAAACAAAGAAAGAGCCAAAGACGAGCTCAGGAAAGG 2138
QY 657 AspLysGlnTriLeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeu 676
DB 2139 GACAAAGCTGGCTGGAGCATGTGCAGCAGGAGGAGCATCAGAGATCAGAGAACCAAGAACTC 2198
QY 677 HisGluGluLysLysLysArgGluSerValLysLysLysAspGlyGluGluLys 696
DB 2199 CACGAAGAGGAAAACTGAAAAGGAGAGAGTGTCAAAAGAGAGATGGCGAGGAAAAA 2258
QY 697 GlyLysGlnGluAlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAla 716
DB 2259 GGCAAAACAGAGACACAAAGCAAGCTGGGTGGCTTTTCCATCAACCAAGAACCAAGCT 2318
QY 717 LysProAlaValGlnAlaProTyrSerThrAlaGluLysGlyProLeuThrIleSerAla 736
DB 2319 AAGCCAGCTGTCCAGCACCCTGGTCCACTGCAGAAAAAAGGTCCACTTACCATTTCTGCA 2378
QY 737 GlnGluAsnValLysValValTyrArgAlaLeuTyrProPheGluSerArgSerHis 756
DB 2379 CAGGAAAAATGTAAAGTGTGTATACCGGCGCACTGTACCCCTTTGAAATCCAGAACCAT 2438
QY 757 AspGluIleThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGlu 776
DB 2439 GATGAATCACTATCCAGCAGGAGACATAGTCATGTTGATGAAAGCAAACTGGAGAA 2498
QY 777 ProGlyTyrLeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaIleThrAla 796
DB 2499 CCGCGCTGCTGGAGGAGAAATTAAGGAAAGACAGGGTGGTTCCCTGCAACTATGCA 2558
QY 797 GluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThrSer 816
DB 2559 GAGAAAAATCCAGAAAAATGAGGTTCCCGCTCCAGTGAAACCAAGTGAATCAATCAATCT 2618
QY 817 AlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSer 836
DB 2619 GCCCTGCCCCCAAACTGGCCCTTGGTGAGACCCCGCCCTTTGGCAGTAACCTCTTCA 2678
QY 837 GluProSerThrThrProAsnTyrAlaAspPheSerSerThrTyrProThrSerThr 856
DB 2679 GAGCCCTCCACGACCCCTAATAACTGGCCGCACTTCAGCTCCACGTCGCGCCACAGCAG 2738
QY 857 AsnGluLysProGluThrAspAsnTyrAspAlaTyrAlaAlaGlnProSerLeuThrVal 876
DB 2739 AATGAGAAACCAAGAACGGAATTAACCTGGGATGATGGGAGCCAGCCCTCTCACCGTT 2798
QY 877 ProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySer 896
DB 2799 CCAAGTGGCGGCGAGTTAAGGAGAGGTCGCGCTTTACTCCAGCCACGCGCACTGGCTCC 2858
QY 897 SerProSerProValLeuGlyGlnGlyLysValGluGlyLeuGlnAlaLeu 916
DB 2859 TCCCCGTCTCTGTGTAGCCGAGCTGACTTCCCTC-----CTC 2897
QY 917 TyrProTyrArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnValIleThr 936
DB 2898 CACCCC----- 2903
QY 937 ValLeuGluGlnGlnAspMetTrpTrpPheGlyGluValGlnGlyGlnLysGlyTrpPhe 956
DB 2904 -----TCCATGAGACTAGGCCACATG 2924
QY 957 ProLysSerTyrValLysLeuIleSerGlyProIleArgLysSerThrSer----- 973
DB 2925 CAGCCCCGGATTGTCTCTTCCCGGACCCCGCTTCCACTCCAGCTCGAGGCTTTTG 2984
QY 974 -----MetAspSerGlySerSerGluSerPro 982
DB 2985 CCCATGCTCGGCCCTCGGCTCGAGTGCCTTCTCCGAGCCCTTCCAGTGTCAAGTCT 3044
QY 983 AlaSerLeuLysArgValAlaSerProAlaLysProVal 996


```
Db      3045  TCTCATCTTCAGGCCCATCTCAGATGCTGCCCTCTGTG 3086
RESULT 10
AAI63825
ID  AAI63825 standard; cDNA; 3466 BP.
XX
AC  AAI63825;
XX
DT  22-OCT-2001 (first entry)
XX
DE  Human polynucleotide SEQ ID NO 33.
XX
KW  Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW  cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW  fungicide; ophthalmologic; cytostatic; immunosuppressive; nootropic;
KW  neuroprotective; anti-allergic; hepatotropic; antidiabetic;
KW  anti-inflammatory; antitumor; anticonvulsant; antibacterial;
KW  antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW  cardiovascular disorder; neurological disease; infection; human; ss.
XX
OS  Homo sapiens.
XX
PN  W0200155308-A2.
XX
PD  02-AUG-2001.
XX
PF  17-JAN-2001; 2001WO-US001309.
XX
XX  31-JAN-2000; 2000US-0179065P.
XX  04-FEB-2000; 2000US-0180628P.
XX  24-FEB-2000; 2000US-0184664P.
XX  02-MAR-2000; 2000US-0186350P.
XX  16-MAR-2000; 2000US-0189874P.
XX  17-MAR-2000; 2000US-0190076P.
XX  18-APR-2000; 2000US-0198123P.
XX  19-MAY-2000; 2000US-0205151P.
XX  07-JUN-2000; 2000US-0209467P.
XX  28-JUN-2000; 2000US-0214886P.
XX  30-JUN-2000; 2000US-0215135P.
XX  07-JUL-2000; 2000US-0216647P.
XX  07-JUL-2000; 2000US-0216880P.
XX  11-JUL-2000; 2000US-0217487P.
XX  11-JUL-2000; 2000US-0217496P.
XX  14-JUL-2000; 2000US-0218290P.
XX  26-JUL-2000; 2000US-0220963P.
XX  26-JUL-2000; 2000US-0220964P.
XX  14-AUG-2000; 2000US-0224518P.
XX  14-AUG-2000; 2000US-0224519P.
XX  14-AUG-2000; 2000US-0225213P.
XX  14-AUG-2000; 2000US-0225214P.
XX  14-AUG-2000; 2000US-0225266P.
XX  14-AUG-2000; 2000US-0225267P.
XX  14-AUG-2000; 2000US-0225268P.
XX  14-AUG-2000; 2000US-0225270P.
XX  14-AUG-2000; 2000US-0225447P.
XX  14-AUG-2000; 2000US-0225757P.
XX  14-AUG-2000; 2000US-0225758P.
XX  14-AUG-2000; 2000US-0225759P.
XX  18-AUG-2000; 2000US-0226279P.
XX  22-AUG-2000; 2000US-0226681P.
XX  22-AUG-2000; 2000US-0226686P.
XX  22-AUG-2000; 2000US-0227182P.
XX  23-AUG-2000; 2000US-0227009P.
XX  30-AUG-2000; 2000US-0228924P.
XX  01-SEP-2000; 2000US-0229287P.
XX  01-SEP-2000; 2000US-0229343P.
XX  01-SEP-2000; 2000US-0229344P.
XX  01-SEP-2000; 2000US-0229345P.
XX  05-SEP-2000; 2000US-0229503P.
XX  05-SEP-2000; 2000US-0229513P.
XX  06-SEP-2000; 2000US-0230437P.
XX  06-SEP-2000; 2000US-0230438P.
XX  08-SEP-2000; 2000US-0231242P.
XX  08-SEP-2000; 2000US-0231243P.
XX  08-SEP-2000; 2000US-0231244P.
XX  08-SEP-2000; 2000US-0231413P.
XX  08-SEP-2000; 2000US-0231414P.
XX  08-SEP-2000; 2000US-0231415P.
XX  08-SEP-2000; 2000US-0232080P.
XX  08-SEP-2000; 2000US-0232081P.
XX  12-SEP-2000; 2000US-0231968P.
XX  14-SEP-2000; 2000US-0232397P.
XX  14-SEP-2000; 2000US-0232398P.
XX  14-SEP-2000; 2000US-0232399P.
XX  14-SEP-2000; 2000US-0232400P.
XX  14-SEP-2000; 2000US-0232401P.
XX  14-SEP-2000; 2000US-0233063P.
XX  14-SEP-2000; 2000US-0233064P.
XX  14-SEP-2000; 2000US-0233065P.
XX  21-SEP-2000; 2000US-0234223P.
XX  21-SEP-2000; 2000US-0234274P.
XX  25-SEP-2000; 2000US-0234997P.
XX  25-SEP-2000; 2000US-0234998P.
XX  26-SEP-2000; 2000US-0235484P.
XX  27-SEP-2000; 2000US-0235834P.
XX  27-SEP-2000; 2000US-0235836P.
XX  29-SEP-2000; 2000US-0236327P.
XX  29-SEP-2000; 2000US-0236367P.
XX  29-SEP-2000; 2000US-0236368P.
XX  29-SEP-2000; 2000US-0236369P.
XX  29-SEP-2000; 2000US-0236370P.
XX  02-OCT-2000; 2000US-0236802P.
XX  02-OCT-2000; 2000US-0237037P.
XX  02-OCT-2000; 2000US-0237038P.
XX  02-OCT-2000; 2000US-0237039P.
XX  13-OCT-2000; 2000US-0237040P.
XX  13-OCT-2000; 2000US-0239335P.
XX  13-OCT-2000; 2000US-0239337P.
XX  20-OCT-2000; 2000US-0240960P.
XX  20-OCT-2000; 2000US-0241221P.
XX  20-OCT-2000; 2000US-0241785P.
XX  20-OCT-2000; 2000US-0241786P.
XX  20-OCT-2000; 2000US-0241787P.
XX  20-OCT-2000; 2000US-0241808P.
XX  20-OCT-2000; 2000US-0241809P.
XX  20-OCT-2000; 2000US-0241826P.
XX  01-NOV-2000; 2000US-0244617P.
XX  08-NOV-2000; 2000US-0246474P.
XX  08-NOV-2000; 2000US-0246475P.
XX  08-NOV-2000; 2000US-0246476P.
XX  08-NOV-2000; 2000US-0246477P.
XX  08-NOV-2000; 2000US-0246523P.
XX  08-NOV-2000; 2000US-0246524P.
XX  08-NOV-2000; 2000US-0246525P.
XX  08-NOV-2000; 2000US-0246526P.
XX  08-NOV-2000; 2000US-0246527P.
XX  08-NOV-2000; 2000US-0246528P.
XX  08-NOV-2000; 2000US-0246532P.
XX  08-NOV-2000; 2000US-0246609P.
XX  08-NOV-2000; 2000US-0246610P.
XX  08-NOV-2000; 2000US-0246611P.
XX  08-NOV-2000; 2000US-0246613P.
XX  17-NOV-2000; 2000US-0249207P.
XX  17-NOV-2000; 2000US-0249208P.
XX  17-NOV-2000; 2000US-0249209P.
XX  17-NOV-2000; 2000US-0249210P.
XX  17-NOV-2000; 2000US-0249211P.
XX  17-NOV-2000; 2000US-0249212P.
XX  17-NOV-2000; 2000US-0249213P.
XX  17-NOV-2000; 2000US-0249214P.
XX  17-NOV-2000; 2000US-0249215P.
XX  17-NOV-2000; 2000US-0249216P.
XX  17-NOV-2000; 2000US-0249217P.
XX  17-NOV-2000; 2000US-0249218P.
XX  17-NOV-2000; 2000US-0249244P.
```

PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX PI
 XX Rosen CA, Barash SC, Ruben SM;
 XX P-PSDB; AAM43519.
 DR WPI: 2001-488781/53.
 DR P-PSDB; AAM43519.
 XX
 PT New isolated nucleic acids and polypeptides, useful for diagnosing,
 PT treating and/or preventing human diseases and disorders.
 XX
 PS Claim 1; SEQ ID NO 33; 664pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAM63803-AAI64012) and
 CC the encoded proteins (AAM434497-AAM43660) useful for preventing, treating
 CC or ameliorating medical conditions e.g. by protein or gene therapy. The
 CC genes were isolated from a range of human tissues disclosed in the
 CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
 CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
 CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
 CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
 CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
 CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3466 BP; 1056 A; 829 C; 861 G; 710 T; 0 U; 10 Other;
 Alignment Scores:
 Pred. No.: 1,456-235 Length: 3466
 Score: 4612.00 Matches: 921
 Percent Similarity: 91.72% Conservative: 9
 Best Local Similarity: 90.83% Mismatches: 28
 Query Match: 78.24% Indels: 56
 DB: 4 Gaps: 4
 US-09-720-934-2 (1-1143) x AAI63825 (1-3466)
 QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
 DB 277 ATGGCTCAGTTTCCACACGCTTTGGTGGCAGCGCTGATATCGGCCATAACTGTAGAG 336
 QY 21 GluArgAlaIleHisAspGlnGlnPheHisSerLeuIysProIleSerGlyPheIleThr 40
 DB 337 GAAAGAGCGAAGCATGATCAGCAGTTCCTCATAGTTTAAGCCAAATATCTGGATTCTACT 396
 QY 41 GlyAspGlnAlaArgAsnPhePhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 DB 397 GGTGATCAAGCTAGAACTTTTCTTCTCATCTGGTTTACCTCACTCTGTTTAGCACAG 456

QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 DB 457 ATATGGGCACTAGCTGACATGAATATGATGGAGAATGATGATCAAGTGGAGTTTCATA 516
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 DB 517 GCTATGAAACTTATCAAACTGAAGCTACAAGGATATCAGCTACCTCTCGCACTTCCCCCT 576
 QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyLe 120
 DB 577 GTCATGAAACAGCAACAGCTGCTATTCTTAGGGCACCAAGCATTTGGTATGGGAGGTATC 636
 QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 DB 637 GCCAGCATGCCACCGCTTACAGCTGTGCTCCAGTGCATGGATGCAATCCAGTTGTT 696
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 DB 697 GGAATGTCTCCAACTAGTATCTTCTGTCCACAGCAGCTGTGCCCTCCCTGGCTAAC 756
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 DB 757 GGGGCTCCCTCTGTATACAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
 QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 DB 817 AACAGTTCTTCTTTAGTAGATCTGCTCCAGGGTCACTAACTAACTAACTAACTAACTAA 876
 QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
 DB 877 GCACAGCTCATTTGATGTGGCCAGTCCACACAGTGGCAGTGGGTGCTTCTTCAGTCA 936
 QY 221 SerArgLeuLysTrpArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 DB 937 TCAAGACTGAAATACAGGCAATTTATTCATAGTCATGACAAACTATGAGTGGACACTTA 996
 QY 241 Thr-----GlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGln 256
 DB 997 ACAGGTTCTGTTTAGTTCCTCCACAGCAAGACTATTTCTATGAGCTCAAGTTTACCACAG 1056
 QY 257 AlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAla 276
 DB 1057 GCTCAGCTGGCTCAATATGGAATCTTCTGACATTCATGATCAAGATGGAACCTTACAGCA 1116
 QY 277 GluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuPro 296
 DB 1117 GAGGAATTTATCTGGCAATGACCTCATTTGATGATGCTATGCTGGCCAACTCCACTGSCA 1176
 QY 297 ProValLeuProGluTyrIleProProSerPheArgValArgSerGlySerGly 316
 DB 1177 CCTGCTCTGCTCCAGATACATTCACCTCTTTTAGAAGAGTTCGATCTGGCAGTGT 1236
 QY 317 IleSerValIleSerSerThrSerValAspGlnArgLeuProGluGluProValLeuGlu 336
 DB 1237 ATATCTGTCATAGCTCAACATCTGTAGATCAGAGGTACACAGAGAACCTAGTTTAGAA 1296
 QY 337 AspGluGlnGlnGlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGlu 356
 DB 1297 GATGAAACAACAATTAGAAAAGAAATTTACTCTGTAACGTTTGAAGATGAAGCGGGAG 1356
 QY 357 AsnPheGluArgGlyAsnLeuGluLeuLysArgArgGlnAlaLeuLeuGluGlnGln 376
 DB 1357 AACTTTGAACGTGGCAACCTGGAACCTGGAGAACGAGGCAAGCTCTCTGGAACAGCAG 1416
 QY 377 ArgLysGluGlnGlnArgLeuAlaGlnLeuGluArgAlaGluGlnGlnArgLysGluArg 396
 DB 1417 CGCAAGGAGCAGGAGCGCTGGCCAGCTGGAGCGCGGAGCAGCAGGAGGAGGAGCGCT 1476
 QY 397 GluArgGlnGlnGlnArgLysArgGlnLeuGluLeuGluLysGlnLeuGlnLysGln 416
 DB 1477 GAGGCCAGGAGCAGGAGCGCAAAAGCAACTGGAACTGGAGAACTGGAAGCAACTGGAAGCAG 1536

Db 1382 ATTTCAGGGCCCATAGGAAGTCTACAGCATGGATTCTGGTTCTTTCAGAGATCCTGCT 1441
 QY 984 SerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGluGluLeuAla 1003
 Db 1442 AGTCTAAAGCGAGTAGCTCTCCAGCAGCCAGCGGTGGTTCGGGAGAGAAATGGCC 1501
 QY 1004 GlnValLeuAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGln 1023
 Db 1502 CAGGTATTGCTTCATACACCGCACCGGCCCGGAGAGCTCACTCTCGCCCTGGTTCAG 1561
 QY 1024 LeuLeuLeuLeuArgLysLysAsnProGlyGlyTyrTrpGluGlyGluLeuGlnAlaArg 1043
 Db 1562 CTGATTTTGATCCGAAAAAGAACCCAGGTGGATGGTGGGAGGAGAGCTGCAAGCACGT 1621
 QY 1044 GlyLysLysArgGlnLeuGlyTyrPheProAlaAsnTyrValLysLeuLeuSerProGly 1063
 Db 1622 GGGAAAAAGCGCCAGATAGGTGGTTCCTCCAGCTAAATTATGTAAAGCTTCTAAGCCCTGGG 1681
 QY 1064 ThrSerLysLeuThrProThrGluProLysSerThrAlaLeuAlaAlaValCysGln 1083
 Db 1682 ACAGCAAAATCACTCCACAGAGCCCACTTAAGTCAACAGCATATTGGCGCAGTGGCCAG 1741
 QY 1084 ValLeuGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAlaPheAsnLysGly 1103
 Db 1742 GTGATTGGGATGTACGACTACCCCGCAGAAATGACGATGAGCTGGGCTTCAACAAGGCGC 1801
 QY 1104 GlnLeuLeuAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGluValAsnGly 1123
 Db 1802 CAGATCATCAACGCTCTCAACAAGGAGGACCCCTGACTGGTGGGAAGGAGAAAGTCAATGGA 1861
 QY 1124 GlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
 Db 1862 CAGTGGGGCTCTTCCCATCAATATGTGAAGCTGACCACAGACATGACCCCAAGCCAG 1921

RESULT 12
 AAZ34573
 ID AAZ34573 standard; cDNA; 2079 BP.
 AC AAZ34573;
 XX
 XX 01-FEB-2000 (first entry)
 DT
 XX Human SH3D1A cDNA clone 5.
 DE
 XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 136..2079
 FT CDS /*tag= a
 FT
 XX WO9953062-A2.
 XX
 XX 21-OCT-1999.
 PD
 XX 16-APR-1999; 99WO-US008371.
 PF
 XX 16-APR-1999; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA
 XX
 XX Korenberg JR, Chen X;
 PI
 XX WPI; 1999-633829/54.
 DR
 XX P-PSDB; AAY32157.
 DR
 XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.

XX Claim 2; Fig 12; 99pp; English.
 PS
 XX This is the nucleotide sequence of a cDNA clone, termed clone 5,
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 2079 BP; 700 A; 451 C; 495 G; 433 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,46e-163 Length: 2079
 Score: 3246.00 Matches: 645
 Percent Similarity: 100.00% Conservative: 3
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 55.06% Indels: 0
 DB: 2 Gaps: 0

US-09-720-934-2 (1-1143) x AAZ34573 (1-2079)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTyrAlaIleThrValGlu 20
 Db 136 ATGGCTCAGTTTCCACACCTTTTGGTGGCAGCTGGGATATCTGGCCATAAATGTAGAG 195
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 196 GAAAGAGCGCAAGCATGATCAGCAGTTCATATGTTAAAGCCATATCTGGATTCAATTA 255
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 256 GGTGATCAAGCTAGAAACTTTTTTTTCAATCTGGTTACTCACTCACTGTTTACACAG 315
 QY 61 IleTyrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 316 ATATGGGCACCTAGCTGACATGAATAATCATGGAAGATGGATCAAGTGGAGTTTCCATA 375
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerSerAlaLeuPro 100
 Db 376 GCTATGAAACTTATCAAACTGAAGCTACAAGATATCAGTACCTCTGCACTTCCCTT 435
 QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 Db 436 GTCATGAAACAGCAACCAAGTTCATATTTCTAGCGCACCCAGCATTTGGTATGGAGTATC 495
 QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 496 GCCAGCATGCCACCGCTTACAGCTGTGCTCCAGTGCCTCAATGGGATGCCAGTATC 555
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 Db 556 GGAATGTCTCCAAACCTTAGTATCTTCTGTTCCACAGCAGCTGTGCCCCCTTGGCTAAC 615
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180

Db 616 GGGGCTCCCTGTTATATACAACTCTGCTGCAATTTGCTCATCTCTGAGCCACATTTGCCA 675
QY 181 LysSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db 676 AAGAGTTCTTCCTTTAGTAGATCTGGTCCAGGGTCACAACTAAACACTAAATTTACAAAG 735
QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
Db 736 GCACAGTCATTTGATGTGGCCAGTGTCCACCACTGCGAGTGGGCTGTTCCTCAGTCA 795
QY 221 SerArgLeuLysTrpArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db 796 TCAGACTGAATACAGCCAAATATTCAATAGTCATGACAAACTATGAGTGGACACTTA 855
QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
Db 856 ACAGGTCCCCAAGCAAGAACTATTCTTATGTCAGTCAAGTTTACCACAGGCTCAGCTGGCT 915
QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
Db 916 TCAATATGGAATCTTTCTGACATTTGATCAAGATGGAAACTTACAGCAGAGGAATTTATC 975
QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProValLeuPro 300
Db 976 CTGGCAATGCACTCATTTGATGTAGCTATGCTGGCCCACTGCCCACCTGCTCTGCTGCT 1035
QY 301 ProGluTrpIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle 320
Db 1036 CCAGATAACATCCACCTCTTTTAGAAGATTCGATCTGGCAGTGGTATATCTGTGCATA 1095
QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
Db 1096 AGCTCAACATCTGTAGATCAGAGGCTACCAGAGAACCACTTTTAGAAGATGAACACAA 1155
QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysArgGluAsnPheGluArg 360
Db 1156 CAATTTAGAAAGAAATTTACTGTGAAGCTTTGAAGATAAGAGCGGAGAACTTTTGAACGT 1215
QY 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
Db 1216 GGCAACTGGAACTGGAGAAACCAAGCAAGCTCTCTCGAAGCAGCAGCGCAGAGGAGCAG 1275
QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu 400
Db 1276 GAGCGCTGGCCAGCTGGAGCGGGCGGAGCAGAGAGGAGGAGCGTGGAGCGCCAGGAG 1335
QY 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
Db 1336 CAAGACGGCAAGACCACTGGAACTGGAGAACCACTGGAAAGCAGCGGAGGAGCTAGAA 1395
QY 421 ArgGlnArgGluGluArgArgLysGluIleGluArgArgGluAlaAlaLysArgGlu 440
Db 1396 CGGCAGAGAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1455
QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgGlnGlnLeuLeuLeuAsnGln 460
Db 1456 CTTGAAGGCACGACCACTTGTAGTGGGACCGAAATCGAAGGCAAGCACTACTAAATCAA 1515
QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
Db 1516 AGAAACAAGAACAGAGGACATAGTTGTACTGAAAGCAAGAAAGAACTTTGGAATTT 1575
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
Db 1576 GAATTTAGAAGCTCTAAATGATAAAAGCATCACTAGAAGGGAACCTTCAAGATATCAGA 1635
QY 501 CysArgLeuThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg 520
Db 1636 TGTGCAATTGACCAACCCCAAGGCAAGAAATTTAGAGGACCAACAAATCTAGAGAGTTGAGA 1695
QY 521 IleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArg 540
Db 1696 ATTGCCGAATACCCATCTACAGCAACAATTCAGGAATCTCAGCAAAATGCTTGAAGA 1755

QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1756 CTTATTCCAGAAAAACAGATACTCAATGACCAATTAATAACAAGTTTCAGCAGAACAGTTTG 1815
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1816 CACAGAGATTCACTTGTTCACCTTAAAGAGCCCTTAGAAGCAAAAGAACTAGCTCGGCAG 1875
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
Db 1876 CACCTACGAGACCACCTGGATGAAGTGGAGAAAGAACTAGATCAAAACTACAGGAGATT 1935
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
Db 1936 GATATTTTCAATAATCAGCTGAAGGAACTAAGAGAAATACACAAATGAAGCAACTCCAG 1995
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGlnArgLysIleIle 640
Db 1996 AAGCAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAGAAACGAAAGATCATA 2055
QY 641 GluLeuGluLysGlnLysGluGlu 648
Db 2056 GAATTAGAAAAAATAAAAAA 2079

RESULT 13

AAK94139
ID AAK94139 standard; cDNA; 2131 BP.
XX
AC AAK94139;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2646.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;
Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
WPI; 2001-524255/58.
P-PSDB; AAK93229.

830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation.

Claim 8; SEQ ID NO 2646; 1380pp + Sequence Listing; English.

The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5' and 3'-ends of the cDNA molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly

CC from EPO

XX Sequence 2131 BP; 694 A; 469 C; 526 G; 442 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,32e-161 Length: 2131
 Score: 3209.00 Matches: 639
 Percent Similarity: 99.53% Conservative: 0
 Best Local Similarity: 99.53% Mismatches: 3
 Query Match: 54.44% Indels: 0
 DB: Gaps: 4

US-09-720-934-2 (1-1143) x AAK941139 (1-2131)

QY 1 MetAlaGlnPheProThrPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
 Db ATGGCTCAGTTCCTCAACACCTTTTGGTGGCAGCCTGGATATCTGGGCCATAACTGTAGAG 264
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 265 GAAAGACGAGCATGATCAGCATCTCCATAGTTTAAAGCCATATCTGGATTCATTACT 324
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 325 GGTGATCAGCTAGAACTTTTTCATCTGGTTACCTCAACCTGTTTATAGCACAG 384
 QY 61 IleTrpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 385 ATATGGGCACCTAGCATGAATGAATGATGAAGAATGGATCAAGTGGAGTTTTCATATA 444
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 445 GCTATGAACCTTATCAAACTGAAGCTACAGGATATCAGTACCTCTGCACTTCCCCCT 504
 QY 101 ValMetLysGlnGlnProValAlaLysSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 Db 505 GTCATGAACACGACCACTTTGCTATTTCTAGCGCCACGACCATTTGGTATGGAGGTATC 564
 QY 121 AlaSerMetProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 565 GCCAGCATGCCACCCCTTACAGCTGTGCTCCAGTGGCAATGGGATCCATTCCTCGATGTT 624
 QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProLeuAlaAsn 160
 Db 625 GGAATGTCTCCAACTAGTATCTCTGTTCCACAGCAGCTGTGCCCCCTGGCTAAC 684
 QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 Db 685 GGGGCTCCCCCTGTTATACAACTCTGCTGCAATTTGCTATCTCTGACGCCACATTTGCCA 744
 QY 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 Db 745 AAGAGTTCCTCTTTAGTAGATCTGCTCAGGTGCTCAGGTGCTCAGGTGCTCAGGTGCT 804
 QY 201 AlaGlnSerPheAspValAlaSerValProValAlaGluTrpAlaValProGlnSer 220
 Db 805 GCACAGCTATTGATGTGCCAGTGTCCACAGTGGCAGAGTGGCTGTTCTTCAGTCA 864
 QY 221 SerArgLeuLysArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 Db 865 TCAAGACTGAATACAGCAATTAATCAATAGTCTATGACAAACTATGAGTGGACACTTA 924
 QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
 Db 925 ACAGTGTCCCAAGCAAGAACTATCTATGCAAGTCAAGTTTACCACAGGCTCAGCTGGCT 984
 QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
 Db 985 TCAATATGAATCTTCTGACATTTGATCAAGATGGAATACTTACAGCAGAGGAATTTATC 1044
 QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
 Db 1045 CTGGCAATGCACCTCATTGATGTAGTATGTCTGGCCAACTGCTGCCACCTGTCTGCTGCT 1104

QY 301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
 Db 1105 CCAGATACATTCACCTCTTTTAGAAGAGTTCGATCTGCAGTGTATATCTGTGCTATA 1164
 QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
 Db 1165 AGCTCAACATCTCTAGTCTCAGAGCTACACAGAGGAACCACTTTTAGAAGATGACAAACA 1224
 QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
 Db 1225 CAATTGAAAGAAATTTACTCTGTAACTTTGAAGATGAAGACGGGAGAACTTTGAACGT 1284
 QY 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
 Db 1285 GGCACCTCGAAGCTGGAGAAACGAAGCAAGCTCTCTCGAACAGCAGCGCAAGAGCAG 1344
 QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGlnArgLysGluArgGluArgGlnGlu 400
 Db 1345 GAGCGCTGGCCAGCTGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1404
 QY 401 GlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
 Db 1405 CAAGAGCGCAAAAGACAACCTGGAACTGGAGAGCACTGGAAAGCAGCGGAGCTAGAA 1464
 QY 421 ArgGlnArgGluGluGluArgArgLysGluIleGluArgGluAlaAlaLysArgGlu 440
 Db 1465 CGGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1524
 QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGluLeuLeuAsnGln 460
 Db 1525 CTTGAAAGGCCAACACCACTTTGAGTGGGAACGGAATCGAAGGCAAGAACTACTAATCA 1584
 QY 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThrLeuGluPhe 480
 Db 1585 AGAAACAAAGAACAAAGAGGACATAGTTGTACTGAAAGCAAAAGAAAAGACATTTGAA 1644
 QY 481 GluLeuGluAlaLeuAsnAspLysHisGlnLeuGluGluGluLysLeuGlnAspIleArg 500
 Db 1645 GAATTAGAGCTCTTAATGATAAAAGCANTCAACTAGAAAGGGAACCTTCAAGATATCAGA 1704
 QY 501 CysArgLeuThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArg 520
 Db 1705 TGTGATTCAGCACCACCAAGGCAAGAAATGAGAGCAAAACAAATCTAGAGATTGAGA 1764
 QY 521 IleAlaGluIleThrHisLeuGlnGlnLeuGlnGluSerGlnGlnMetLeuGlyArg 540
 Db 1765 ATTGCCGAATTCACCCATCTACAGCAACAATTACAGGAATCTCAGCAAAATGCTTGAAGA 1824
 QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
 Db 1825 CTTATTCCAGAAACACAGATCTCAATGACCAATTTAAACCAAGTTTCAGCAGAACAGTTG 1884
 QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
 Db 1885 CACAGAGATTCACTGTTTACACTTAAAGAGCCTTAGAAGCAAAAGAACTAGATCAAACTACAGAGATT 1944
 QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIle 600
 Db 1945 CACCTACAGACCACTGGATGAGTGGAGAAAGAACTAGATCAAACTACAGAGATT 2004
 QY 601 AspIlePheAsnAsnGlnLeuLysGluIleArgGluIleHisAsnLysGlnGlnLeuGln 620
 Db 2005 GATATTTTCAATAATCAGCTGAAGAACTTAAGAGAAATACACAAATGAAGCAACCTCCAG 2064
 QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGluGluArgLysIleIle 640
 Db 2065 AAGCAAAAGTCCATGGAGCTGACACTGAAACAGAAAGAAACAGAAAGAAATGATATA 2124
 QY 641 GluLeu 642
 Db 2125 GAATTA 2130

RESULT 14
AAS84762
ID AAS84762 standard; cDNA; 2874 BP.
XX AAS84762;
AC
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #20566.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG20575.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 20566; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2874 BP; 940 A; 679 C; 703 G; 552 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 3,29e-157 Length: 2874
Score: 3133.50 Matches: 625
Percent Similarity: 84.34% Conservative: 5
Best Local Similarity: 83.67% Mismatches: 12
Query Match: 53.16% Indels: 105
DB: 5 Gaps: 6

US-09-720-934-2 (1-1143) x AAS84762 (1-2874)

QY 436 AlaAlaLysArgGluLeuGluAArgGlnArgGlnLeuGluTrpGluArgAsnArgGln 455

Db 1 GCTGCAAAACGGGAACTTCANAGGCAACGACCACTTGAGTGGGAACGGAAATCGAAGCGAA 60
QY
456 GluLeuLeuAsnGlnArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLys 475
Db 61 GAACCTACTAAATCAAGAAACAAAGAACAAAGAGGACATAGTTGCTGAAAGCAAGAA 120
QY 476 LysThrLeuGluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLys 495
Db 121 AAGACTTTGGAAATTTGAATTAGAAAGCTTAATGATAAAAGCATCAACTAGAAAGGAAA 180
QY 496 LeuGlnAspIleArgCysArgLeuThrGlnArgGlnGluLeuGluSerThrAsnLys 515
Db 181 CTTCAAGATATCAGATGTGATTGACCAACCCAAAGCAAGAAATCAGAGCACAAACAA 240
QY 516 SerArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnGluSerGln 535
Db 241 TCTAGAGAGTTGAGAAATTCGGGAATCACCATCTACAGCAACAAATACAGGAATCTCAG 300
QY 536 GlnMetLeuGlyArgLeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnVal 555
Db 301 CAAATGCTTTGGAAGACTTATTCAGAAAAACAGATACTCAATGACCAATTTAAACCAAGTT 360
QY 556 GlnGlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLys 575
Db 361 CAGCAGAACAGTTTGCACAGAGATTCACTTGTGTACACTTTAAAGAGCCCTTAGAAGCAAAA 420
QY 576 GluLeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSer 595
Db 421 GAACCTAGCTCGGAGCAGCCCTACGAGACCACCACTGGATGAGTGGAGAAAGAACTAGATCA 480
QY 596 LysLeuGlnGluIleAspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsn 615
Db 481 AAACCTACAGAGATTGATATTTTCAATAATCAGCTGAAGGAACTAAGAGAAATACACAAT 540
QY 616 LysGlnGlnLeuGlnLysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGln 635
Db 541 AAGCAACAATTCAGAAAGCAAAAGTCCATGGAGGCTGAACGACTGAACAGAAAGAACAA 600
QY 636 GluArgLysIleLeuGluLeuGlnLysGlnLysGluGluAlaGlnArgAlaGlnGlu 655
Db 601 GAACGAAAGATCTAGATTTAGAAAAACAAAGAAAGAGAGCCCAAGACGAGCTCAGGAA 660
QY 656 ArgAspLysGlnTrpLeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLys 675
Db 661 AGGCAACAAGCAGTGGCTGGAGCATGTGCAGCAGGAGCAGCAGCATCAGAGCAACCAAGAAA 720
QY 676 LeuHisGlnGluLysLysLysArgGluGluSerValLysLysLysAspGlyGluGlu 695
Db 721 CTCACGAGAGAGAAAAACTGAAAAAGGAGGAGAGTGTCAAAAGAAAGGATGGCGAGGAA 780
QY 696 LysGlyLysGlnGluAlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluPro 715
Db 781 AAGGCAACAGCAAGCAACAGCAAGCTGGTTCGGCTTTCCATCAACACCAAGAACCA 840
QY 716 AlaLysProAlaValGlnAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSer 735
Db 841 GCTAAGCAGCTGTCCAGGCACCTGGTCCACTGCAGAAAAAGAGTCCACTTACCAATTTCT 900
QY 736 AlaGlnGluAsnValLysValValTyrArgAlaLeuTyrProPheGluSerArgSer 755
Db 901 GCACGAAAAATGTAAGTGGTGTATTACCGGGCACTGTACCCCTTTTGAATCCAGAAGC 960
QY 756 HisAspGluIleThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGly 775
Db 961 CATGATGAATCACTATCCAGCCAGAGACATAGTCTGTTGGTGAAGCAACCACTGGA 1020
QY 776 GluProGlyTrpLeuGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyr 795
Db 1021 GAACCCGGCTGGCTTGGAGAGAAATTTAAAGGAAAGACAGGGGTGGTTCCTCGCAACTAT 1080
QY 796 AlaGluLysIleProGluAsnGluValProAlaProValLysProValThrAspSerThr 815

Db 1081 GCAGAGAAATCCAGAAATAGGTTCCCGCTCCAGTGAACAGGAGCTGATTCAACA 1140
QY 816 SerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSer 835
Db 1141 TCTGCCCTGCCCGCAACCTGCGTGGAGACCGCCCGCTTTGGCAGTAACCTCT 1200
QY 836 SerGluProSerThrThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSer 855
Db 1201 TCAGAGCCCTCCAGACCCCTTAATCTGGGCGGACTTCAGTCCAGTGGCCACCAGC 1260
QY 856 ThrAsnGluLysProGluThrAspAsnTrpAlaAlaTrpAlaAlaGlnProSerLeuThr 875
Db 1261 ACGNATCAGAAACCCAGAAACGATACTGGGATGATGGGCGGAGCCAGCCCTCTCTCACC 1320
QY 876 ValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGly 895
Db 1321 GTTCCAGTGGCGGCCAGTTAAGGCAGAGGTCCGCTTACTCCAGCCAGCGCCACTGGC 1380
QY 896 SerSerProSerProValLeuGlyGlnGlyValGluGlyLeuGlnAlaGlnAla 915
Db 1381 TCCTCCCGCTCTCTGTGTAGGCGAGGTGAAGGTGGAGGGCTACAGCTCAAGCC 1440
QY 916 LeuTyrProTyrArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIle 935
Db 1441 CTATATCTTGGAGAGCCAAAGACACACCTTTAAATTTTAAACAAATGATGTCATC 1500
QY 936 ThrValLeuGluGlnGlnAspMetTrpTrpPheGlyGluValGln----- 950
Db 1501 ACCGTCTGGAACAGCAAGCATGTGTGGTGTGGAGAGTTCAAGAAATTTATTGCCATG 1560
QY 950 ----- 950
Db 1561 TACACTTACGAGAGTCTGACAGAGGAGATTAACTTTCAGCAAGGGGATGCTATTG 1620
QY 951 -----GlyGlnLys---GlyTrpPhe 956
Db 1621 GTTACCAAGAAAGATGTGACTGTGGAGAGAACAGTGGCGGCAAGGCGGAGTCTTC 1680
QY 957 ProLysSerTyrValLysLeuLeuSerGlyProLeuArgLysSerThrSerMetAspSer 976
Db 1681 CTTTCTAATATGTGAGCTT-----AAGATTTACAGGGGTCT 1719
QY 977 GlySerSerGluSerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProVal 996
Db 1720 GGAAGTCTGGGAAACAGGAGTTTAGGAAA-----AACTCT--- 1758
QY 997 ValSerGlyGluGluAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGln 1016
Db 1759 -----GAAATGGCCAGGTTATTGCTCATACCGCCAGCCGCGCCGAGCAG 1806
QY 1017 LeuThrLeuAlaProGlyGlnLeuIleLeuIleArgLysLysAsnProGlyGlyTrpTrp 1036
Db 1807 CTCACCTCGCCCTGGTGCAGTCTGATTTGATTCGAAAGAAAGAACCCCA----- 1854
QY 1037 GluGlyGluLeuGlnAlaArgLysLysLysArgGlnIleGlyTrpPheProAlaAsnTyr 1056
Db 1854 ----- 1854
QY 1057 ValLysLeuLeuSerProGlyThrSerLysIleThrProThrGluProProLysSerThr 1076
Db 1854 ----- 1854
QY 1077 AlaLeuAlaAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAsp 1096
Db 1855 -----GTGTGCCAGTGTGATGGATGTACGACTACACCGCCGCAAGTACGAT 1902
QY 1097 GluLeuAlaPheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTrp 1116
Db 1903 GAGCTGGCTTCAACAGGCGCAGATCATCAAGCTCTCAACAGGAGGACCTGACTGG 1962
QY 1117 TrpLysGlyGluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThr 1136
Db 1963 TCGAAAGGAGAGTCAATGGACAGTGGGGCTCTTCCCATCCAAATTATGTGAAGCTGACC 2022

QY 1137 ThrAspMetAspProSerGln 1143
Db 2023 ACAGACATGACCCAGCCAG 2043
RESULT 15
AAH16578
ID AAH16578 standard; cDNA; 2131 BP.
XX AAH16578;
AC AAH16578;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:15658.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EPI074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 15658; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of oligonucleotides comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 2131 BP; 691 A; 448 C; 572 G; 420 T; 0 U; 0 Other;

Alignment Scores: 2.79e-148 Length: 2131
Pred. No.: 2962.50 Matches: 594
Score:

Percent Similarity:	94.15%	Conservative:	1
Best Local Similarity:	93.99%	Mismatches:	0
Query Match:	50.25%	Indels:	37
DB:	4	Gaps:	1
US-09-720-934-2 (1-1143) x AAH16578 (1-2131)			
QY	1	MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleThrValGlu	20
Db	347	ATGGCTCAGTTTCCACACACTTTTGGTGGCAGCTGGATATCTGGGCCATAACTGTAGAG	406
QY	21	GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProlleSerGlyPheIleThr	40
Db	407	GAAGAGCGAAGCATGATCAGCAGTTCCATAGTTTAAAGCCAATATCTGGATTCAATTACT	466
QY	41	GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln	60
Db	467	GGTGATCAAGCTAGAAACTTTTTTTTCAATCTGGGTACCTCAACCTGTTTTAGCACAG	526
QY	61	IleTrrAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle	80
Db	527	ATATGGGCACTAGCTACATGATATATGATGGAGATGATCAAGTGGAGTTTTCATATA	586
QY	81	AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro	100
Db	587	GCTATGAACTTATCAAACTGGAAGCTACAAAGGATATCAGCTACCTCTGCACCTTCCCT	646
QY	101	ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle	120
Db	647	GTCAATGAAACAGCAACAGTTGCTATTTCTAGCGCACCACTTT-----	691
QY	121	AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal	140
Db	691	-----	691
QY	141	GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn	160
Db	692	-----GCAGCTGTGCCCTCCCTGGCTTAAC	715
QY	161	GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro	180
Db	716	GGGCTCTCCCTGTTATACAACTCTGCTGCAATTTGCTCATCTCGCCGACCAATTGCCA	775
QY	181	LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys	200
Db	776	AAGAGTTCTTCTTTAGTAGATCTGGTCCAGGTCACACTAAACACTAAATTACAAAAG	835
QY	201	AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer	220
Db	836	GCACAGTCATTTGATGTGGCCAGTGTCCACACAGTCGAGAGTGGGCTGTTCTCTCAGTCA	895
QY	221	SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu	240
Db	896	TCAAGACTGAAATACAGGCAATATTTCATAGTATGATGATGATGATGATGATGATGATG	955
QY	241	ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla	260
Db	956	ACAGTCTCCCAAGCAAGAACTATTCTTATGAGTCAAGTTTACCACAGGCTCAGCTGGCT	1015
QY	261	SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle	280
Db	1016	TCAATATGGAATCTTCTGACATTCATGATCAAGATGGAATACTTACAGCAGAGGAATTTATC	1075
QY	281	LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro	300
Db	1076	CTGGCAATGACCTCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1135
QY	301	ProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle	320
Db	1136	COAGATACATCTCCCTCTTTTGAAGAGTTCGATCTGGCAGTGGTATATCTGTCATA	1195
QY	321	SerSerThrSerValAspGlnArgLeuProGluProValLeuGluAspGlnGln	340

Search completed: August 3, 2004, 04:38:52
Job time : 1081 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 3, 2004, 03:20:38 ; Search time 196 seconds

(without alignments)

3236.270 Million cell updates/sec

Title: US-09-720-934-2

Perfect score: 5895

Sequence: 1 MAQFPFFGSLDIWAVE.....QVGLFSPYVYKLTMDPSQ 1143

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09720934/runat_29072004_164338_1244/app_query_fasta_1.1287
-DB=Issued_Patents_NA -QPMT=fastap -SUFFIX=jul29.rn1 -MINWATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934 @CGN 1.1 105 @runat_29072004_164338_1244 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1728.5	29.3	2017	4	US-09-404-879A-72
2	1728.5	29.3	2017	4	US-09-338-933-72
3	1728.5	29.3	2017	4	US-09-215-681-72
4	1728.5	29.3	2017	4	US-09-216-003A-72
5	1089.5	18.5	2873	4	US-08-630-915A-193
6	634	10.8	747	4	US-08-630-915A-193
7	477.5	8.1	4165	1	US-08-095-737-1
8	477.5	8.1	4165	1	US-08-480-145-1
9	477.5	8.1	4165	2	US-08-477-389-1
10	467	7.9	531	4	US-09-404-879A-5
11	467	7.9	531	4	US-09-338-933-5
12	467	7.9	531	4	US-09-215-681-5

13	467	7.9	531	4	US-09-216-003A-5	Sequence 5, Appli
14	465	7.9	3033	1	US-08-095-737-3	Sequence 3, Appli
15	465	7.9	3033	1	US-08-480-145-3	Sequence 3, Appli
16	465	7.9	3033	2	US-08-477-389-3	Sequence 3, Appli
17	456	7.7	480	4	US-09-404-879A-60	Sequence 60, Appl
18	456	7.7	480	4	US-09-338-933-60	Sequence 60, Appl
19	456	7.7	480	4	US-09-215-681-60	Sequence 60, Appl
20	456	7.7	480	4	US-09-216-003A-60	Sequence 60, Appl
21	434	7.4	9551	1	US-08-056-200-93	Sequence 93, Appl
22	434	7.4	9551	2	US-08-800-644-93	Sequence 93, Appl
23	389.5	6.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
24	389.5	6.6	3489	4	US-09-298-568-1	Sequence 1, Appli
25	389.5	6.6	3489	4	US-09-410-399-1	Sequence 1, Appli
26	389.5	6.6	32207	2	US-08-770-379-20	Sequence 20, Appl
27	389.5	6.6	32207	3	US-08-757-669A-20	Sequence 20, Appl
28	389.5	6.6	32207	4	US-09-230-371A-20	Sequence 20, Appl
29	373	6.3	531	4	US-09-404-879A-4	Sequence 4, Appli
30	373	6.3	531	4	US-09-338-933-4	Sequence 4, Appli
31	373	6.3	531	4	US-09-215-681-4	Sequence 4, Appli
32	373	6.3	531	4	US-09-216-003A-4	Sequence 4, Appli
33	354	6.0	4248	4	US-10-164-595-53	Sequence 53, Appl
34	335	5.7	4056	4	US-10-164-595-55	Sequence 55, Appl
35	333.5	5.7	1659	1	US-08-475-894-1	Sequence 1, Appli
36	333.5	5.7	1659	1	US-08-484-710-1	Sequence 1, Appli
37	333.5	5.7	1659	2	US-08-484-709-1	Sequence 1, Appli
38	333.5	5.7	1659	3	US-08-474-697-1	Sequence 1, Appli
39	333.5	5.7	3143	4	US-08-671-354-1	Sequence 1, Appli
40	333.5	5.7	3222	4	US-08-714-741-39	Sequence 39, Appl
41	326	5.5	6644	4	US-08-875-435B-5	Sequence 5, Appli
42	318.5	5.4	5325	4	US-10-164-595-17	Sequence 17, Appl
43	317	5.4	5334	4	US-10-164-595-21	Sequence 21, Appl
44	314.5	5.3	3543	4	US-09-543-681A-2264	Sequence 2264, Ap
45	309	5.2	1414	4	US-09-023-655-1356	Sequence 1356, Ap

ALIGNMENTS

RESULT 1

US-09-404-879A-72
; Sequence 72, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-72

Alignment Scores:			
Pred. No.:	1.6e-123	Length:	2017
Score:	1728.50	Matches:	382
Percent Similarity:	70.63%	Conservative:	99
Best Local Similarity:	56.09%	Mismatches:	141
Query Match:	29.32%	Indels:	59
DB:	4	Gaps:	18

US-09-720-934-2 (1-1143) x US-09-404-879A-72 (1-2017)

Qy	1	MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu	20
Db	42	ATGGCTCAGTTTCCACAGCGATCAATGGAGGCCAAATATGTGGCTATTACATCTGAA	101
Qy	21	GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr	40

Db 102 GAACGTACTAGCATGATAACAGTTTCATTAACCTCAACCTTCAGGAGGTTCATACAA 161
Qy
41 GlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 162 GGTGATCAAGCCCGTACTTTTCTACAGTCAGTCTGCGCGCCCGGTTCAGTGA 221
Qy 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGGCTTATCAGATCTGAACAGATGGAGATGGACCAAGAGTTCCTATATA 281
Qy 81 AlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAACACTCATCAAGTAAAGTTGAGGCGCCACACAGCTGCTGTAGTCTCCCTCT 341
Qy 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAACAAACCCCTATGTCTCTCCACTAATCTCTGCT---CGTTTGGGATGGA 398
Qy 119 GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal 132
Db 399 -----AGCATGCCCAATCTGTCCATTCATCAGCCATTCCTCCAGTTGCACCTATA 449
Qy 133 -----ProMetGly-----SerIleProValValGlyMetSerPro 144
Db 450 GCAACACCTGTCTCTGCTACTTCAGGAGCACGATTCCTCCCTAATGATGCTGCT 509
Qy 145 ThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 510 CCCCTAGTGCCTCTCTGTAGTACATCTCTCATTTACCA-----AATGGAAGTCCCACT 560
Qy 165 ValIleGlnProLeuProAlaPheAlaIleProAlaAlaThrLeuProLysSerSerSer 184
Db 561 CTCATTCAGCCTTTA-----TCCATTCCTTATCTTCTTCAACATTCCTCCATGCACTCA 617
Qy 185 PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGln 202
Db 618 TACAGCCTGATGATGGGAGGATTTGTT-----GGTGTAGTATCCAGAGGCCCGAC 668
Qy 203 Ser-----PheAspValAlaSerValProPro----- 211
Db 669 TCTCTGATGATTTAGGATCTAGTAGCTCAACTCTCCCTCAACTGCTTCCCTCTCAGGGAAC 728
Qy 212 -----ValAlaGlnTrpAlaValProGlnSerSerArgLeuLysTyrArg 226
Db 729 TCACCTAAGACAGGACCTCAGAGTGGGAGTCTCTCAGCCTTCAAGATTAAGTATCG 788
Qy 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 789 CAAAAATTTAATAGTCTAGACAAAGGCGATGAGCGGATACCTCTCAGGTTTTCAGCTAGA 848
Qy 247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 849 AATGCCCTTCTTCACTCAATCTCTCAAACTCAGCTAGTACTATTGGGACTCTGGCT 908
Qy 267 AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMethIleLeu 286
Db 909 GACATCGATGGTGCAGCAGAGTTGAAGCTGAAGATTTATTTCTGGCGATGACCTCACT 968
Qy 287 AspValAlaMetSerGlyGlnProLeuProProValLeuProProGluTyrIleProPro 306
Db 969 GACATGGCCAAAGCTGGACAGCCATACCACTGAGCTGTGCTCCCTCCGAGCTTGTCCCTCA 1028
Qy 307 SerPheArgArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326
Db 1029 TCTTTT-----AGAGGGGGAAGCAAGTTGATTCGTGTAATGGAAC----- 1070
Qy 327 GlnArgLeuProGluProValLeuGluAspGluGlnGlnGlnLeuLysLysLeu 346
Db 1071 -----CTGCTTTCATATCAGAAAACACAGAGAGAGAGCCTCAG-----AAGAACTG 1118
Qy 347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGlyAsnLeuLeuGlu 366

Db 1119 CCAGTTACTTTTGGAGGACAAACGGAAAGCCCAACTATGAAACGAGGAAACATGAGCTGGAG 1178
Qy 367 LysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGlnGluArgLeuAlaGlnLeu 386
Db 1179 AAGCGACGCCAAGTGTGTATGGAGCAGCAGCAGGAGGAGGCTGAACCGAAGACCCAGAA 1238
Qy 387 GluArgAlaGluGlnGlnArgLysGluArgGlnGlnGlnGlnGlnGlnArgLysArgGln 406
Db 1239 GAGAAGGAGAGTGGAGCGGACCAACAGAGAGAACTGCAAGAGAGAAATGGAAGAGCAG 1298
Qy 407 LeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGluArgGlnGlnGluGlu 426
Db 1299 CTGAGTTGGAGAAACGCTTGGAGAAACAGAGAGAGCTGGAGAGACAGCGGAGGAGAAAG 1358
Qy 427 ArgArgLysGluIleGluArgGluAlaAlaLysArgGluLeuGluArgGlnArgGln 446
Db 1359 AGGAGAAAGAGAGTAGAAGACGAGAGCAGCAAAACAGGAGCTTTGAGAGACACCGCGT 1418
Qy 447 LeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGlnArgAsnLysGluGlnGlu 466
Db 1419 TTAGATGGGAAAGACTCCGTCGCGAGGAGCTGCTCAGTCAGAAGACCAGGAAACAGAA 1478
Qy 467 AspIleValValLeuLysAlaLysLysThrLeuGluPheGluLeuGluAlaLeuAsn 486
Db 1479 GACATTTGTCAGGCTGAGCTCCAGAAAGAAAGTCTCCACCTGGAACTCGAAGCAGTGAAT 1538
Qy 487 AspLysLysHisGlnLeuGluGluLysLeuGlnAspIleArgCysArgLeuThrThrGln 506
Db 1539 GGAAAAACATCAGCAGATCTCAGGAGACTACAGATGTCCAAATCAGAAAGCAACACAA 1598
Qy 507 ArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHis 526
Db 1599 AAGACTGAGCTAGAAGTTTGGATAAACACAGTGTGACCTGGAAATTTATGAAATCAACAA 1658
Qy 527 LeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArgLeuIleProGluLysGln 546
Db 1659 CTTCAACAGAGCTTAAGGAATATCAAAATAAGCTTATCTATCTGCTCCCTGAGAGCAG 1718
Qy 547 IleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuVal 566
Db 1719 CTATTAAACGAAGAAATTAACAAACATGCGCTCAGTACACACA---CCTGATTCAGGGATC 1775
Qy 567 ThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAsp 584
Db 1776 AGTTTACTTCAATAAAGTCAACAAAAGGAAGAAATATGTC---CAAAGACTTAAGAA 1832
Qy 585 GlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIleAspIlePheAsn 604
Db 1833 CAATTAGATGCTCTTGAAGAAAGAAACTGCATCTAAGCTCTCAGAAATGGAATTCATTITAC 1892
Qy 605 AsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLysGlnLysSer 624
Db 1893 AATCAGCTGAAGAACTCAGAGAAAGCTATAATACACAGCAGTTAGCCCTTGAACAATT 1952
Qy 625 MetGluAlaGluArgLysGlnLysGluGlnGluArgLysLysIleLeuLeuGluLys 644
Db 1953 CATAAATCAAAACGTCACAAATTTGAAGGAATTCGAAGAAAGAAAGATTAGACAAAGAAA 2012
Qy 645 Gln 645
Db 2013 AAA 2015

RESULT 2

US-09-338-933-72
; Sequence 72, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933

```
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-72

Alignment Scores:
Pred. No.: 1.6e-123      Length: 2017
Score: 1728.50          Matches: 382
Percent Similarity: 70.63%      Conservative: 99
Best Local Similarity: 56.09%    Mismatches: 141
Query Match: 29.32%             Indels: 59
DB: 4                         Gaps: 18

US-09-720-934-2 (1-1143) x US-09-338-933-72 (1-2017)

QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleTrpAlaIleThrValGlu 20
DB 42 ATGGCTCAGTTTCCACAGCGATGAATGAGGGCCAAATATGTGGCTATTACATCTGAA 101

QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
DB 102 GAACGTACTAAGCATGATAACAGTTTGATAACCTCAACCTTCAGGAGGTTACATAACA 161

QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
DB 162 GGTGATCAAGCCGTACTTTTCTTACAGTCAGGTCTGCGGCCCGGTTTAGCTGAA 221

QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
DB 222 ATATGGCCCTTATCAGATCTCGAAAGGATGGGAAGATGACCAAGAGTTCTCTATA 281

QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
DB 282 GCTATGAACCTATCAAGTTAAAGTTGCGAGGCCAACAGCTGCTGTAGTCCCTCCT 341

QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
DB 342 ATCATGAAACAACCCCTATGTCTCTCCACTAATCTCTGCT---CGTTTGGGATGGGA 398

QY 119 GlyIleAlaSerMetPro-----AGCATGCCCAATCTGTCCATTCATCAGCCATTGCTCCAGTTGCACCTATA 449
DB 399 -----ProMetGly-----SerIleProValValGlyMetSerPro 144

QY 133 -----ProMetGly-----SerIleProValValGlyMetSerPro 144
DB 450 GCACACACCTTGTCTCTGCTACTTCAGGACCAAGTATTCCTCCCTTAATGATGCTGCT 509

QY 145 ThrLeuValSerSerValProThrAlaAlaValProLeuAlaAsnGlyAlaProPro 164
DB 510 CCCCTAGTGCCTCTCTGTAGTACATCTCTATTACCA-----AATGGAATGCCAGT 560

QY 165 ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSer 184
DB 561 CTCATTTCAGCCCTTA---TCCATTCCTTATCTCTTCAACATTGCTCATGCTCATCT 617

QY 185 PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGln 202
DB 618 TACAGCTGATGATGGAGGATTTGGT-----GGTGCTAGTATCCAGAGGCCAG 668

QY 203 Ser-----PheAspValAlaSerValProPro----- 211
DB 669 TCTCTGATTGATTAGGATCTAGTAGCTCAACTTCCTCAACTGCTTCCCTCTCAGGGAAC 728

QY 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysTyrArg 226
DB 729 TCACCTTAAGACAGGGACCTCAGAGTGGGAGTTCTCTCAGCCTTCAAGATTAAAGTATCG 788

QY 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
DB 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
```

789 CAAAAATTTAATAGTCTTAGACAAAGGCATGAGCGGATACCTCTCAGGTTTTTCAAGTAGA 848

247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266

849 AATGCCCTTCTTCTAGTCAAAATCTCTCAAACTCAGCTAGCTACTATTGGACTCTGGCT 908

267 AspIleAspGlnAspGlyLysLeuThrAlaGluPheIleLeuAlaMetHisLeuIle 286

909 GACATCGATGCTGACGACAGATTGAAAGCTGAAGATTTATTCTGGCGATGCACCTCACT 968

287 AspValAlaMetSerGlyGlnProLeuProValLeuProGlnProGlyLysIleProPro 306

969 GACATGCCCAAGCTGGACAGCCACTACCACTGACGTGTGCTCCGAGCTTGTCCCTCCA 1028

307 SerPheArgArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326

1029 TCTTTC-----AGAGGGGGAAGCAAGTTGATTCTGTTAATGGAATCT----- 1070

327 GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnLeuGlnLysLysLeu 346

1071 -----CTGCTTTCATATCAGAAAAACACAAGAAAGAGAGCTCAG-----AAGAACTG 1118

347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGlyAsnLeuLeuGlu 366

1119 CCAGTTACTTTTGGAGCAACCGAAAGCCCACTATGAACGAGAAACATGGAGCTGGAG 1178

367 LysArgArgGlnAlaLeuLeuGluGlnArgLysGluGlnGlnArgLeuAlaGlnLeu 386

1179 AAGCGACGCCAGTGTGTGAGGACGACGACGAGGAGGCTGAACCAAGGCCGAGAA 1238

387 GluArgAlaGluGlnGluArgLysGluArgGluGlnGlnGluArgLysArgGln 406

1239 GAGAAGAGAGTGGGAGCGGAAACAGAGAGAACTGCAAGAGCAAGAAATCGAAGAACGAG 1298

407 LeuGluLeuLysGlnLeuGluLysGlnArgGluLeuGluArgGlnArgGluGlu 426

1299 CTGGAGTTGGAGAAACGCTTGGAGAAACAGAGAGAGCTGGAGAGACAGCGGAGGAGAG 1358

427 ArgArgLysLeuIleGluArgArgGluAlaLysArgGluLeuGluArgGlnArgGln 446

1359 AGGAGAAAGGAGATAGAAAGACGAGAGGCGAGCAAAACAGGAGCTTGAGAACACGCCGT 1418

447 LeuGluTrpIleArgAsnArgArgGlnGlnLeuLeuAsnGlnArgAsnLysGluGln 466

1419 TTAGAATGGAAAGACTCCGTCGGCAGGAGCTGCTCAGTCAGACAGACCAGGAGAACAA 1478

467 AspIleValValLeuLysAlaLysLysLysThrLeuGluPheGluLeuAlaLeuAsn 486

1479 GACATTTGTGAGGCTGAGCTCCAGAAAGAAAGTCTCCACCTGGAACTGGAAAGCAGTGAAT 1538

487 AspLysLysHisGlnLeuGluLysLeuGlnAspIleArgCysArgLeuThrThrGln 506

1539 GGAAACATCAGCAGATCTCAGCGACACTACAGATGTCCAAATCAGAAGCAACACAA 1598

507 ArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHis 526

1599 AAGACTGACTAGAAAGTTTTGGATAAAAGTGTGACCTGGAAATTTATGCAATCAAACAA 1658

527 LeuGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArgLeuIleProGluLysGln 546

1659 CTTCAACAGAGCTTAAGGAATATCAAAATAAGCTTATCTATCTGCTGCTCCTGAGAGCAG 1718

547 IleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuVal 566

1719 CTATTAAACGAAGAAATTAACATGCTAGCTCAGTAAACACA---CCTGATTCAGGATC 1775

567 ThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAsp 584

1776 AGTTTACTTCAATAAAAGTCAACAGAAAGGAAGAAATATATGC---CAAAAGACTTAAAGAA 1832

585 GlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIleAspIlePheAsn 604

1833 CAATTAGTGTCTCTTGAAGAAAGAACTGCATCTAAGCTCTCAGAAATGATTCATTAAAC 1892

QY	605	AsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGlnLysSer	624
Db	1893	AATCAGCTGAAGAACTCAGAGAAAGCTATAATACACAGCAGTAGCCCTTGAACTT	1952
QY	625	MetGluAlaGluArgLeuLysGlnLysGluGluArgLysIleIleGluLeuLys	644
Db	1953	CATAAATCAACAGCTGACAAATTCAGGAATCGAAGAAATCGAAGAAAGATAGAGCAAAAAA	2012
QY	645	Gln 645	
Db	2013	AAA 2015	
RESULT 3			
US-09-215-681-72			
; Sequence 72, Application US/09215681A			
; Patent No. 6528253			
; GENERAL INFORMATION:			
; APPLICANT: Mitcham, Jennifer L.			
; APPLICANT: Frudakis, Tony N.			
; APPLICANT: King, Gordon E.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS			
; TITLE OF INVENTION: OF OVARIAN CANCER			
; FILE REFERENCE: 210121.463			
; CURRENT APPLICATION NUMBER: US/09/215.681A			
; CURRENT FILING DATE: 1998-12-17			
; NUMBER OF SEQ ID NOS: 310			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 72			
; LENGTH: 2017			
; TYPE: DNA			
; ORGANISM: Homo sapien			
US-09-215-681-72			
Alignment Scores:			
Pred. No.:	1-66-123	Length:	2017
Score:	1728.50	Matches:	382
Percent Similarity:	70.63%	Conservative:	99
Best Local Similarity:	56.09%	Mismatches:	141
Query Match:	29.32%	Indels:	59
DB:	4	Gaps:	18
US-09-720-934-2 (1-1143) x US-09-215-681-72 (1-2017)			
QY	1	MetAlaGlnPheProThrPheGlySerLeuAspIleTyrAlaIleThrValGlu	20
Db	42	ATGGCTCAGTTTCCACAGCGATGAATGGAGGGCCAAATATGTGGGCTATTACATCTGAA	101
QY	21	GluArgAlaLysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr	40
Db	102	GAACGTACTAAGCATGATACAGTTTGTATACCTCAACCTTCAGGAGGTTACATAACA	161
QY	41	GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln	60
Db	162	GCTGATCAAGCCGCTACTTTTCTCAGCTCAGGCTGCGCGCCCGGTTTAGCTGAA	221
QY	61	IleTyrAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle	80
Db	222	ATATGGGCTTATCAGATCTGAACAGGATGGAGATGGACAGCAGAGTTCCTATA	281
QY	81	AlaMetLysLeuIleLysLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro	100
Db	282	GCTATGAACCTCATCAAGTTAAAGTTGAGGCCCAACAGCTGCTGTAGTCTCCTCTCT	341
QY	101	ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly	118
Db	342	ATCATGAACAACCCCTTATGTCTCTCCACTAATCTCTGCT---CGTTTGGGATGGGA	398
QY	119	GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal	132
Db	399	-----AGCATGCCCAATCTGCTCATTATCAGCAATTCCTCCAGCTTGACCTATA	449
QY	133	-----ProMetGly-----SerIleProValValGlyMetSerPro	144

Db	450	GCAACACCCCTTGTCTTCTGCTACTTCCAGGACAGTAGTATTCCTCCCTAAATGATGCTGCT	509
QY	145	ThrLeuValSerSerValProThrAlaAlaValProLeuAlaAsnGlyAlaProPro	164
Db	510	CCCTTAGTGCCTTCTGTGTAGTACATCTCTCAITACCA-----AATGAGCTGCCAGT	560
QY	165	ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSer	184
Db	561	CTCATTCAGCTTTA---TCCATTCTTATTTCTTCTTCAACATTGCTCTCATGATCATCT	617
QY	185	PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuLysAlaGln	202
Db	618	TACAGCTGATGATGGAGGATTGGT-----GGTGTAGTATCCAGAGGCCAG	668
QY	203	Ser---PheAspValAlaSerValProPro-----	211
Db	669	TCCTGATTGATTTAGGATCTTAGTAGCTCACTTCTCAACTGCTTCTCTCAGGGAAC	728
QY	212	-----ValAlaGluTyrAlaValProGlnSerSerArgLeuLysTyrArg	226
Db	729	TCACCTAAGACAGGACCTCAGAGTGGCAGTTCCTCAGCCTTCAAGATTAAGTATCGG	788
QY	227	GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg	246
Db	789	CMAAAATTTAATAGTCTAGACAAAGGATGAGCGGATACCTCTCAGGTTTTCAGGCTAGA	848
QY	247	ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSer	266
Db	849	AATGCCCTTCTTCAGTCAATCTCTCAAACTCAGCTAGTACTATTGGACTGCTGCT	908
QY	267	AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle	286
Db	909	GACATCGATGCTGAGCGGACAGTTGAAAGCTCAAGAAATTTATCTGGCGATGCACCTACT	968
QY	287	AspValAlaMetSerGlyGlnProLeuProValLeuProGluTyrIleProPro	306
Db	969	GACATGCCAAAGCTGGACAGCCACTACCATGCTGCTGCTCCGAGCTTGTCTCTCCA	1028
QY	307	SerPheArgArgValArgSerGlySerGlyIleSerValIleSerSerThrValAsp	326
Db	1029	TCCTTTC-----AGAGGGGGAAACCAAGTTGATTCTGTATGGAAGT-----	1070
QY	327	GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnGlnLeuGluLysLysLeu	346
Db	1071	-----CTGCCTTCATATATCAGAAACACACAGAGAGAGCCTCAG-----AAGAACTG	1118
QY	347	ProValThrPheGluAspLysArgGluAsnPheGluArgGlyAsnLeuGluLeuGlu	366
Db	1119	CCAGTTACTTTTGAGGACAAACCGGAAAGCCAACTATGACGAGGAAACATGAGAGCTGGAG	1178
QY	367	LysArgArgGlnAlaLeuGluGlnGlnArgLysGluGlnGluArgLeuAlaGlnLeu	386
Db	1179	AAGCAGCCCAAGCTTGTGATGGAGCAGCAGCAGAGGAGGCTGAACCGCAAGCCAGAAA	1238
QY	387	GluArgAlaGluGlnGluArgLysGluArgGluGlnGlnGlnGluArgLysArgGln	406
Db	1239	GAGAAAGGAGAGTGGGAGCGGAAACAGAGAGAACTGCAAGACCAAGATGGAAGAGCAG	1298
QY	407	LeuGluLeuLysGlnLeuLysGlnArgGluLeuGluArgGlnArgGluGluGlu	426
Db	1299	CTGGAGTTGGAGAAACGCTTGGAGAAACAGAGAGCTGGAGAGCAGCGGAGGAGAGAG	1358
QY	427	ArgArgLysGluIleGluArgArgGluAlaLysArgGluLeuGluArgGlnArgGln	446
Db	1359	AGCAGAAAGGAGATAGAAAGACAGAGGCGAGGAGCAAAACAGAGCTTGAGAGACAACGCGT	1418
QY	447	LeuGluTyrGluArgAsnArgArgGlnGluLeuLeuAsnGlnArgAsnLysGlnGlnGlu	466
Db	1419	TTAGAATGGGAAGACTCCGTCGGAGGAGCTGCTCAGAGACACCGAGGAAACAGAA	1478
QY	467	AspIleValValLeuLysAlaLysLysThrLeuGluPheGluLeuGluAlaLeuAsn	486

```
Db 1479 GACATTCTCAGGCTGAGCTCCAGAAAGAAAGTCTCCACCTGGAACTGGAGCAGTGAAT 1538
QY AsplysHisGlnLeuGluGlyLeuGlnAspIleArgCysArgLeuThrThrGln 506
Db 1539 GGAAGAACATCAGCAGATCTCAGGCAGACTCAAGATGTCACAAATCAGAAAGCAACACAA 1598
QY ArgGlnGluLeuGluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHis 526
Db 1599 AAGACTGAGCTAGAAAGTTTGGATAAAGAGTGTGACCTGGAAATATGGAATCAAAACAA 1658
QY LeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArgLeuIleProGluLysGln 546
Db 1659 CTTCAACAGAGCTTAAGGATATCAAAATAGCTTATCTATCTGGTCCCTGAGAACGAG 1718
QY IleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuVal 566
Db 1719 CTATTAAACCAAGAAATTAATAACATGCAGCTCAGTAACACA--CCTGATTGAGGATC 1775
QY ThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAsp 584
Db 1776 AGTTTACTTCATAAAAGTTCATCAGAAAGAGAGATTATGCT--CAAGACTTAAAGAA 1832
QY GlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIleAspIlePheAsn 604
Db 1833 CAATTAGATGCTCTTGAAGAAAGAACTGCATCTAAGCTCTCAGAAATGGATTCAATTAC 1892
QY AsnGlnLeuLysGlnLeuArgGluIleHisAsnLysGlnGlnLeuGlnLysGlnLysSer 624
Db 1893 AATCAGCTGAAGAACTCAGAGAAAGCTATAATAACACAGCAGTTAGCCCTTGAACACTT 1952
QY MetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIleGluLeuGluLys 644
Db 1953 CATAAATCAACGTCACAAATTAAGGAAATCGAAAGAAAGAAATAGAGCAAAAAA 2012
QY 645 Gln 645
Db 2013 AAA 2015

RESULT 4
US-09-216-003A-72
; Sequence 72, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216.003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-72

Alignment Scores:
Pred. No.: 1,6e-123 Length: 2017
Score: 1728.50 Matches: 382
Percent Similarity: 70.63% Conservative: 99
Best Local Similarity: 56.09% Mismatches: 141
Query Match: 29.32% Indels: 59
DB: 4 Gaps: 18

US-09-720-934-2 (1-1143) x US-09-216-003A-72 (1-2017)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 42 ATGGCTCAGTTTCCACAGCATGAATGGAGGGCCAAATATATGTGGCTATTACATCTGAA 101
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProLysLeuPheIleThr 40
```

```
Db 102 GAACGTACTAAGCATGATAAACAAGTTTGTATAACCTCAAACTTCAGAGGTTACATAACA 161
QY GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 162 GGTGATCAAGCCCGTACTTTTCTTACAGTCAGGTCTGCGGCCCGGTTTGTAGCTGAA 221
QY IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGCCCTTATCAGATCTGAACAAGGATGGGAAGATGACCAAGCAAGATTTCTCTATA 281
QY AlaMetLysIleLeuLysLeuLysLeuGlnGlyTyroGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAACATCATCAAGTTTAAAGTTTGCAGGGCCCAACAGCTGCTGTAGTCTCCCTCT 341
QY ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAAACAACCCCTTAAGTTTCTTCCACTAATCTCTGCT---CGTTTGGATGGGA 398
QY GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal 132
Db 399 -----AGCATGCCCAATCTGTCATTTCATCAGCCATTGCTCCAGTTGCACCTATA 449
QY PrometGly-----SerIleProValValGlyMetSerPro 144
Db 450 GCAACACCCCTGTCTTCTGCTACTTCAGGGACCAAGTATTCTCCCTCAATATGATGCTGCT 509
QY ThrLeuValSerSerValProThrAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 510 CCCTAGTGCTCTGTAGTACATCTCTCAATACCA-----AATGGAATGCCAGT 560
QY ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSerSer 184
Db 561 CTCATTGAGCTTTA---TCCATTCTCTTCTTCTTCAACATTCCTCATCATCATCT 617
QY PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGln 202
Db 618 TACAGCCTGATGATGGAGGATTGGT-----GGTGTAGTATCCAGAGGCCAG 668
QY Ser---PheAspValAlaSerValProPro----- 211
Db 669 TCTCTGATTGATTAGGATCTAGTCTCAACTTCTCAACTGCTCCCTCTCAGGGAAC 728
QY 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysTyArg 226
Db 729 TCACCTAAGACAGGACCTCAGAGTGGCAGATTCTCAGCCTTCAAGATTAAAGTATCG 788
QY GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 789 CAAAAATTTAATAGTCTAGACAAAGGATGAGCGGATACCTCTCAGGTTTTTCAAGCTAGA 848
QY ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 849 AATGCCCTTCTTCAGTCAATCTCTCAAACTCAGCTAGTACTATTGGACTTGGCT 908
QY AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle 286
Db 909 GACATCGATGTTGACGACAGTGTGAAGCTGAAGATTTATTCTGGCAGATGCACCTCACT 968
QY AspValAlaMetSerGlyGlnProLeuProProValLeuProGluTyIleProPro 306
Db 969 GACATGGCCAAAGCTGGACAGCCACTTACCCTGACCTGACCTGCCCTCCAGCTTGTCCCTCA 1028
QY SerPheArgArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326
Db 1029 TCTTTC-----AGAGGGGAAAGCAAGTTGATTCTGTATCGAACT----- 1070
QY GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnGlnLeuGluLysLysLeu 346
Db 1071 -----CTGCCTTCATATCAGAAAAACACAAGAAAGAGAGCTCAG-----AAGAAACTG 1118
QY ProValThrPheGluAspLysLysArgGluAsnPheGluArgGlyAsnLeuGluLeuGlu 366
```

APPLICANT: McCONNELL, Stephen J.

QY 772 SerGlnThrGlyGluProGlyTrrpLeuGlyGlyGluLeuLysGlyLysThrGlyTrrpPhe 791
Db 318 AAAACCGTAGGAGAACCTGGTCTTTATGGCTTTTCAAGGAAATTTTGGCTGTTT 377
QY 792 ProAlaAsnTyrAlaGluLysIleProGluAsnGluValProAlaProValLysProVal 811
Db 378 CCATGCAATTATGTAGAAATAATGCCATCAAGTGA 413
QY 812 ThrAspSerThrSerAlaProAlaProLysLeuAlaLeuArgGluThrProAlaProLeu 831
Db 414 -----AATGAAAGCTGTATCTCCAAAGAGGCTTACTTCTCTCAGTTCCTTA 467
QY 832 AlaValThrSerSerGluProSerThrThrProAsnAsnTrpAlaAspPheSerThr 851
Db 468 TCTGCTACCTCA -----ACTTCTCT 488
QY 852 TrpProThrSerThrAsnGluLysProGluThrAspAsnTrpAspAlaTrpAlaAlaGln 871
Db 489 GAACCACTTTCTTCAATCAACACAGCATCAGTCACTGATTATCAAAAT---GTATCTTTT 545
QY 872 ProSerLeuThrValProSerAlaGlyGlnLeuArgGlnArgSerAlaPheThrProAla 891
Db 546 TCAAACTTAACCTTAATACATCATGCGAG-----AAAAATCAGCCTTCACCTCGACT 599
QY 892 ThrAlaThrGlySerSerProSerProValLeuGlyGlnGlyGluLysValGluGlyLeu 911
Db 600 GTGTCCCTCGA---TCTGTATCACCTTATTCATGCAGCAGGACAAAGTGGTAGAAAACCTTA 656
QY 912 GlnAlaGlnAlaLeuTyrProTrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLys 931
Db 657 AAAGCAGGCGCCCTTTGTTCTGTGACTGCAAGAAAGATACCACTTGAACCTTCTCAAAA 716
QY 932 AsnAspValIleThrValLeuGluGlnGlnAspMetTrpPheGlyGluValGlnGly 951
Db 717 CATGACATTATTACTGTCTGGAGCAGCAAGAAATTTGGTGTGGGGAGGTGCATGA 776
QY 952 GlnLysGlyTrrpPheProLysSerTyrValLysLeuLysSerGlyPro----- 967
Db 777 GGAAGAGGATGGTTTCCCAATCTTATGTCAAGATTCATCTCGGAGTGAAGTAAACGG 836
QY 968 -----IleArgLysSerThrSer----- 973
Db 837 GAAGAACCAAGCTTTGTATGCAGCTGTAAATAAGAAACCTACCTCGCAGCCTATCA 896
QY 973 ----- 973
Db 897 GTTGGAGAAGATATATTGCATTTATCCATATTCAAGTGTGGAACCTGGAGATTGACT 956
QY 974 -----MetAspSerGlySerSerGlu-- 980
Db 957 TTCACAGAAGGTGAAGAAATATTGGTGACCCAGAAAGATGGAGAGTGTGGACAGGAAGT 1016
QY 981 -----SerProAlaSerLeuLysArgVa 988
Db 1017 ATTGGAGATAGAGTGAATTTTCCATCAACTATGTCAACCAAGAGATCAAGAGAT 1076
QY 988 lAlaSerProAlaAlaLysProValValSerGlyGlu-----GluIleAlaGlnValI 1006
Db 1077 TTT-GGGAGTGTAGCAAGCTGGAGCATCAATATAAAACCTGAGATTCTCAGGTAAC 1135
QY 1006 eAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuLe 1026
Db 1136 TTCAGCATATGTGCTCTCTGTCTGTGAACCACTTAGCTTGCACCAAGAGTAAATATT 1195
QY 1026 uIleArgLysLysAsnProGlyGlyTrrpGluGlyGluLeuGlnAlaArgGlyLysLy 1046
Db 1196 AATTCRAAGAAATAACAAGTGGTGGTGGCAGGAGAGTTACAGCCAGAGAAAAA 1255
QY 1046 sArgGlnIleGlyTrrpPheProAlaAsnTyrValLysLeuLeuSerProGlyThrSerLy 1066
Db 1256 GCGACAGAAAGGATGGTTTCTGCCAGCTCATGTATTAACTTTTGGGTCCAAGTAGTGAAG 1315

QY 1066 sIleThrProThrGluProProLysSerThrAlaLeuAlaAlaValCysGlnValIleGl 1086
Db 1316 AGCCACACCTGCGCTTTTCATCT-----GTATGTCAAGTGTGATTGC 1354
QY 1086 yMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPheAsnLysGlyGlnIleI 1106
Db 1355 TATGTATGACTATGCGCAAAATATGAAGATGAGCTCAGTTTCTCCAAAGGCAACTCAT 1414
QY 1106 eAsnValLeuAsnLysGluAspProAspTrpTrpLysGlyGluValAsnGlyGlnValGl 1126
Db 1415 TAATGTATGAACAAGATGATCTGTATGTTGGCAAGGAGAGATCAACGGGTGACTGG 1474
QY 1126 yLeuPheProSerAsnTyrValLysLeuThrThrAspMetAspProSerGln 1143
Db 1475 TCTCTTTCCTCAAACTACGTTAAGATGAGCAGCAGACTCAGATCCAAAGTCAA 1526

RESULT 6
US-08-630-915A-39
; Sequence 39, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOMLIES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 747 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-630-915A-39

Alignment Scores:
Pred. No.: 5-83e-40 Length: 747
Score: 634.00 Matches: 131
Percent Similarity: 60.51% Conservative: 36
Best Local Similarity: 47.46% Mismatches: 79
Query Match: 10.75% Indels: 30
DB: 4 Gaps: 5

US-09-720-934-2 (1-1143) x US-08-630-915A-39 (1-747)

```
QY 868 TrpAlaAlaGlnProSerLeuThrValProSerAlaGlyGlnLeuArgGlnArgSerAla 887
Db 4 TTCGCGCGCGCGTGCACCAAGATCATCTTGGAGTGAAGTAAACGGGAAGACCGAGAA 63
QY 888 PheThrProAlaThrAlaThrGlySerSerProSerProValLeuGlyGlnGlyGluLys 907
Db 64 GCITTTGATGACGCTGTAATAGAAACCTTACCTCGGCAGCTTATTCAGTTGGAGAA-- 120
QY 908 ValGluGlyLeuGlnAlaGlnAlaLeuThrProTrpArgAlaLysLysAspAsnHisLeu 927
Db 121 -----GAATATATTGCACTTATTCATATTCAGTTCAGTTCGGAACCTGGAGATTG 168
QY 928 AsnPheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 947
Db 169 ACTTTCACAGAGGTGAAGAAATATGTTGACCCAGAAAGATGGAGAGTGGTGACAGGA 228
QY 948 GluValGlnGlyGlnLysGlyTrpPheProLysSerTrpValLysLeuIleSerGlyPro 967
Db 229 AGTATTGGAGTAGAAGTGAATTTTCCATCAAACTATGTCAAA----- 273
QY 968 IleArgLysSerThrSerMetAspSerGlySerSerGluSerProAlaSerLeuLysArg 987
Db 274 ---CCAAAGGATCAAGAGATTTGGCAGTCTAGCAGTCTGGAGCATCAATAA--- 327
QY 988 ValAlaSerProAlaAlaLysProValSerGlyGluGluIleAlaGlnValIleAla 1007
Db 328 -----AAACCT-----GAGATTGCTCAGGTAACTTCA 354
QY 1008 SerTrpThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuIleLeu 1027
Db 355 GCATATGTTGCTTCTGTTCTGACAACTTAGCCTTGCCAGGACAGCTTAATATTAATT 414
QY 1028 ArgLysLysAsnProGlyGlyTrpTrpGluGlyGlnLeuGlnAlaArgGlyLysLysArg 1047
Db 415 CTAAGAAAAATACAAAGTGGTGGTGGCAAGAGATTACAGGCCAGAGAAAAAAGCGA 474
QY 1048 GlnIleGlyTrpPheProAlaAsnTrpValLysLeuLeuSerProGlyThrSerLysIle 1067
Db 475 CAGAAAGGATGGTTCTCCAGATCATGTATAACTTTTGGGTCCAGCAGTGAAGAGCC 534
QY 1068 ThrProThrGluProProLysSerThrAlaLeuAlaAlaValCysGlnValIleGlyMet 1087
Db 535 ACACCTGCTTTCATCCT-----GATGTGCGTGTATTCCTATG 573
QY 1088 TyrAspTrpThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGlnIleLeuAsn 1107
Db 574 TATGACTATGCAGCAAAATATGAAGATGAGTTCAGTTCTCCAGGGCAACTCATTAAT 633
QY 1108 ValLeuAsnLysGluAspProAspTrpTrpLysGlyGluValAsnGlyGlnValGlyLeu 1127
Db 634 GTTATGAACAAAGATGATCCTGATTGGTGGCAAGAGAGATCAACGGGGTGTGCTCTC 693
QY 1128 PheProSerAsnTrpValLysLeuThrThrAspMetAspProSerGln 1143
Db 694 TTTCTCTCAAACTACGTTAAGATGACGACAGACTCAGATCCAAAGTCAA 741

RESULT 7
US-08-095-737-1
; Sequence 1, Application US/08095737
; Patent No. 5487979
; GENERAL INFORMATION:
; APPLICANT: DiFiore, Pier P
; APPLICANT: Fazioli, Francesca
; TITLE OF INVENTION: A Substrate for the Epidermal Growth
; TITLE OF INVENTION: Factor Receptor Kinase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: United States of America
; ZIP: 92660
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,737
FILING DATE: 19930722
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4165 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 21...2709
US-08-095-737-1

Alignment Scores:
Pred. No.: 9,2e-27 Length: 4165
Score: 477.50 Matches: 194
Percent Similarity: 35.13% Conservative: 145
Best Local Similarity: 20.10% Mismatches: 285
Query Match: 8.10% Indels: 341
DB: 1 Gaps: 28

US-09-720-934-2 (1-1143) x US-08-095-737-1 (1-4165)
QY 15 TrpAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 34
Db 384 TGGGCTGTAAACACCTGAAGATAAGCCCAATATGATCAATATTATTGATAGTTTAAGCCCA 443
QY 35 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 54
Db 444 GTGATGATTTCTGCTGCTGGTGAAGTGAACACAGTGTGCTCACTCAAGTTACT 503
QY 55 GlnProValLeuAlaGlnIleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 74
Db 504 GTGGATATCTCTGGAGAGATTGGGAGTTGAGTGATATTGACCATGATGGATGCTTGAC 563
QY 75 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 94
Db 564 AGAGATGAGTTTGCAGTTGCCAATGTTTGGTATTAAGTGTGCTGCTGAG----- 611
QY 95 ProSerAlaLeuProProValMetLysGlnGlnProValAlaIleSerSerAlaProAla 114
Db 611 ----- 611
QY 115 PheGlyMetGlyGlyIleAlaSerMetProProLeuThrAlaValAlaProValProMet 134
Db 612 -----AAAGAACCTGTGCCAATG 629
QY 135 GlySerIleProValValGlyMetSerProThrLeuValSerSerValProThrAlaAla 154
Db 630 -----TCCTTGCTCCAGCTTG 647
QY 155 ValProProLeuAlaAsnGlyAlaProProValIleGlnProLeuProAlaPheAlaHis 174
Db 648 GTGCCACCA----- 656
QY 175 ProAlaAlaThrLeuProLysSerSerSerPheSerArgSerGlyProGlySerGlnLeu 194
```

657	Db	AsnThrLysLeuGlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGlu	665	TCTAAGAGA	665	
195	QY	AsnThrLysLeuGlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGlu	214		214	
666	Db	AAAACG	671		671	
215	QY	TrpAlaValProGlnSerSerArgLeuLysTyArgGlnLeuPheAsnSerHisAspLys	234		234	
672	Db	TGGGTTGTATCCCTCGCAGAAAAGCTAAATATGATGAATCTTCTGTGAAAACCTGATAAA	731		731	
235	QY	ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu	254		254	
732	Db	GATATGACGAGATTTGTGTCTGATTGGAGTCCGTGAAATATTTCTTTGAAAACAGAGTTTA	791		791	
255	QY	ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeu	274		274	
792	Db	CCCTCTACCTTACTAGCCCATATATGCTCAATATGCGACAAAGGACTGTGGAAAGCTT	851		851	
275	QY	ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMet	293		293	
852	Db	TCAAAGGATCAGTTTGCCTTGGCTTTTCACTTAATCAGTCAGAAGTTAATCAAGGGCATT	911		911	
294	QY	ProLeuProValLeuProProGluTyIleProProSerPheArgValArgSer	313		313	
912	Db	GATCCTCCTCACGTTCTTACTCTCTGAAATGATTCCACCATCA	962		962	
314	QY	GlySerGlyIleSerValIleSerSerThrSerValAspGlnArgLeuProGluGluPro	333		333	
963	Db	AGTTTCAAAAGAACATCATAGATCAAGTCCCTGTT	998		998	
334	QY	ValLeuGluAspGluGlnGlnGlnLeuGluLysLysLeuProValThrPheGluAspLys	353		353	
999	Db		1016	GCAGATTTCCTGCTATT	1016	
354	QY	LysArgGluAsnPheGluArgGlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeu	373		373	
1017	Db	AAG	1049	GAACATAGATCTCTTAAACAATGAATAAGTT	1049	
374	QY	GluGlnGlnArgLysGluGlnGluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArg	393		393	
1050	Db	GACCTACAGAGGGGAAAG	1088	AATATGTGGAACAGGACCTT	1088	
394	QY	LysGluArgGluArgGlnGlnGlnGluArgLysArgGlnLeuGluLeuLysGlnLeu	413		413	
1089	Db	AAGAGAAAGAAAGTACTATTAAACAGGACAAAGTGAAGTT	1130		1130	
414	QY	GluLysGlnArgGluLeuGluArgGlnArgGluGluArgLysGluIleGluArg	433		433	
1130	Db		1130		1130	
434	QY	ArgGluAlaAlaLysArgGluLeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArg	453		453	
1130	Db		1130		1130	
454	QY	ArgGlnGluLeuLeuAsnGlnArgAsnLysGluGlnGluAspIleValLeuLysAla	473		473	
1131	Db	CAGATCTTCAAGATGAAGTTCAAGGGAGNATACTAATCTGCAAAACCTACAGGCC	1187		1187	
474	QY	LysLysLysThrLeuGluPheGluLeuGluAlaLeuAsnAspLysHisGlnLeuGlu	493		493	
1188	Db	CAGAAACAGCAGGTACAGGAACCTCTTGTGAATCGATGACGAGAAGCCAGCTGGAG	1247		1247	
494	QY	GlyLysLeuGlnAspIleArgCysArgLeuThrThrGlnArgGlnGluIleGluSerThr	513		513	
1248	Db	GAGCAACTCAGGAAGTCAGAAAGAAATGTGCTGAGGAGGCCCACTGATCTCTCTCTG	1307		1307	
514	QY	AsnLysSerArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnLeuGlnGlu	533		533	
1308	Db	AA	1346	GCTGAATTAAGTCTAGGAATCGAGATCTCCACT	1346	
534	QY	SerGlnGlnMetLysArgLeuIleProGluLysGlnIleLeuAsnAspGlnLeuLys	553		553	

QY 832 -----AlaValThrSerSerGluPro----- 838
 Db 2379 GGGAAAGATCCATCAACAAATGGATTCTCTGATCCCTTTAACTGAATGATCCATTT 2438
 QY 839 -----SerThrThrProAsnAsnTrpAlaAspPhe 848
 Db 2439 CAGCCTTCCAGGCAAGCAGTAGCCCAAGAAAAGATCCTGAAATGTTTGTGATCCA 2498
 QY 849 SerSerThrTrpProThrSerThrAsnGluLysProGluThrAspAsnTrpAlaTrp 868
 Db 2499 TTCACTTCTGCTACTACCACTACCAATAAAGAGGCTGATCCAGCAATTTGCCAACTTC 2558
 QY 869 AlaAlaGlnProSer 873
 Db 2559 AGTGCATTATCCCTCT 2573

RESULT 8
 US-08-480-145-1
 ; Sequence 1, Application US/08480145
 ; Patent No. 5717067
 ; GENERAL INFORMATION:
 ; APPLICANT: DiFiore, Pier P
 ; APPLICANT: Fazioli, Francesca
 ; TITLE OF INVENTION: A Substrate for the Epidermal Growth
 ; TITLE OF INVENTION: Factor Receptor Kinase
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Knobbe, Martens, Olson & Bear
 ; STREET: 620 Newport Center Drive, Sixteenth Floor
 ; CITY: Newport Beach
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,145
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/095,737
 ; FILING DATE: 22-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israelsen, Ned A
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET NUMBER: NIH060,001A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 235-8550
 ; TELEFAX: (619) 235-0176
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4165 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 21...2709
 ; US-08-480-145-1

Alignment Scores:
 Pred. No.: 9,2e-27
 Score: 477.50
 Percent Similarity: 35.13%
 Best Local Similarity: 20.10%
 Query Match: 8.10%

Length: 4165
 Matches: 194
 Conservative: 145
 Mismatches: 285
 Indels: 341

DB: 1 Gaps: 28
 US-09-720-934-2 (1-1143) x US-08-480-145-1 (1-4165)
 QY 15 TrpAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 34
 Db 384 TGGGCTGTAACAACTGAAGATAAGGCCAAATATGATCAATATTTGATAGTTAAGGCCA 443
 QY 35 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 54
 Db 444 GTGAATGGATTCTGTCTGTGATAAAGTGAACACCACTGTTGCTCACTCTAAGTTACCT 503
 QY 55 GlnProValLeuAlaGlnIleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 74
 Db 504 GTGGATATCTTGAAGAGTTTGGAGCTTGAGTGATATTGACCATGATGAATGCTTGAC 563
 QY 75 GlnValGluPheSerIleAlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeu 94
 Db 564 AGAGATGAGTTTGCAGTTGCCATGTTTGTGTACTGTGCACCTGGAG----- 611
 QY 95 ProSerAlaLeuProValMetLysGlnGlnProValAlaIleSerSerAlaProAla 114
 Db 611 ----- 611
 QY 115 PheGlyMetGlyGlyIleAlaSerMetProLeuThrAlaValAlaProValProMet 134
 Db 612 -----AAAGAACCTGTGCCAATG 629
 QY 135 GlySerIleProValValGlyMetSerProThrLeuValSerSerValProThrAlaAla 154
 Db 630 -----TCTTGGCCCTCCAGCCCTG 647
 QY 155 ValProProLeuAlaAsnGlyAlaProProValIleGlnProLeuProAlaPheAlaHis 174
 Db 648 GTGCCACCA----- 656
 QY 175 ProAlaIleThrLeuProLysSerSerPheSerArgSerGlyProGlySerGlnLeu 194
 Db 657 -----TCTAAGAGA 665
 QY 195 AsnThrLysLeuGlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGlu 214
 Db 666 AAAACG----- 671
 QY 215 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 234
 Db 672 TGGTTGTATCCCTCGAGAAAAGCTAAATATGATGAAATCTTCTCTGAAAACCTGATAA 731
 QY 235 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 254
 Db 732 GATATGGACGGATTGTCTGGATTGGAGTCCGTGAAATATTTCTTGAACACAGGTTTA 791
 QY 255 ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeu 274
 Db 792 CCTTCTACCTTACTAGCCCATATATGTCATTATGCGACACAAGGACTGTGGAGCTT 851
 QY 275 ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMet---SerGlyGln 293
 Db 852 TCAAGGATCAGTTTGGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 911
 QY 294 ProLeuProProValLeuProGluTyrIleProProSerPheArgValArgSer 313
 Db 912 GATCCTCTCAGCTTCTTACTCTCTGAATGATTCACCATCA-----GACAGGCC 962
 QY 314 GlySerGlyIleSerValIleSerSerThrSerValAspGlnArgLeuProGluGluPro 333
 Db 963 AGTTTACAAAAGACATCATAGATCAAGTCTCTGTT----- 998
 QY 334 ValLeuGluAspGluGlnGlnGlnLeuGluLysLysLeuProValThrPheGluAspLys 353
 Db 999 -----GCAGATTTCTCTGCTATT 1016
 QY 354 LysArgGluAsnPheGluArgGlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeu 373

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,389
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/095,737
 FILING DATE: 22-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Israelsen, Ned A
 REGISTRATION NUMBER: 29,655
 REFERENCE/DOCKET NUMBER: NIH060.001A
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4165 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 21..2709
 US-08-477-389-1

Alignment Scores:
 Pred. No.: 9.2e-27 Length: 4165
 Score: 477.50 Matches: 194
 Percent Similarity: 35.13% Conservative: 145
 Best Local Similarity: 20.10% Mismatches: 285
 Query Match: 8.10% Indels: 341
 DB: 2 Gaps: 28

US-09-720-934-2 (1-1143) x US-08-477-389-1 (1-4165)

```

175 ProAlaAlaThrLeuProLysSerSerSerPheSerArgSerGlyProGlySerGlnLeu 194
    |||:::
657 -----TCTAAGAGA 665
195 AsnThrLysLeuGlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGlu 214
    |||
666 AAAACG----- 671
215 TrpAlaValProGlnSerSerArgLeuLysTrpArgGlnLeuPheAsnSerHisAspLys 234
    |||
672 TGGTTGTATCCCTCGAGAAAGCTPAATATGATGAATCTCTCGAAACAGTATAA 731
235 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrLeuMetGlnSerSerLeu 254
    |||
732 GATATGGACGGATTGTGTCTGGATTGGAGTCGCTGAATATTTCTTGAACACAGGTTA 791
255 ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeu 274
    |||
792 CCTTCTACCTTACTAGCCCATATATGCTATTATCGCACACAAAGGACTGTGGAGGCTT 851
275 ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMet---SerGlyGln 293
    |||
852 TCAAAGGATCAGTTTGGCTTTCCTTATCATGATCAGTGAAGTTAATCAAGGGCATT 911
294 ProLeuProValLeuProGluTrpIleProSerPheArgValArgSer 313
    |||
912 GATCCTCCTCAGGTTCTTACTCTCTGAAATGATTCACCATCA-----GACAGGGCC 962
314 GlySerGlyIleSerValIleSerSerThrSerValAspGlnArgLeuProGluPro 333
    |||
963 AGTTTACAAAAGAACATCATAGGATCAAGTCCTGTT----- 998
334 ValLeuGluAspGluGlnGlnGlnLeuLysLysLeuProValThrPheGluAspLys 353
    |||
999 -----GCAGATTTCCTGCTATT 1016
354 LysArgGluAsnPheGluArgGlyAsnLeuGluLeuGluLysArgGlnAlaLeuLeu 373
    |||
1017 AAG-----GAACTAGATCTCTTAACAAATGAATAGTT 1049
374 GluGlnGlnArgLysGluGlnGlnArgLeuAlaGlnLeuGluAlaGlnGlnArg 393
    |||
1050 GACCTACAGAGGGGAAAG-----ATAATGTGGAACAGAGCCTT 1088
394 LysGluArgGluArgGlnGlnGlnArgLysArgGlnLeuGluLysGlnLeu 413
    |||
1089 AAGAGAGAGGAGAGTACTATTAAACAGAGGCAAGTGAAGTT----- 1130
414 GluLysGlnArgGluLeuGluArgGlnArgGluGluArgLysGluLeuArg 433
    |||
1130 ----- 1130
434 ArgGluAlaAlaLysArgGluLeuGluArgGlnArgGlnArgGlnArgGlnArg 453
    |||
1130 ----- 1130
454 ArgGlnGluLeuLeuAsnGlnArgAsnLysGluGlnGluAspIleValValLeuLysAla 473
    |||
1131 ---CAGGATCTTCAAGATGAAGTTCAAGGGAGATACTAATCTGCAAAACACTACAGGCC 1187
474 LysLysLysThrLeuGluPheGluLeuAlaLeuAsnAspLysLysHisGlnLeuGlu 493
    |||
1188 CAGAAACAGCAGGAGTACAGAACTCTTGAATCTGGATGAGCAGAAACCCAGCTGGAG 1247
494 GlyLysLeuGlnAspIleArgCysArgLeuThrThrGlnArgGlnGluIleGluSerThr 513
    |||
1248 GAGCAACTCAGGAGAGTCAGAAAGAAATGCTCTGAGGGCCCACTGATCTCTCTCTG 1307
514 AsnLysSerArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnGlnGlu 533
    |||
1308 AAA-----GCTGAATTAATAGTCAGGAATCGCATCTCCACT 1346
534 SerGlnGlnMetLeuGlyArgLeuLeuProGluLysGlnIleLeuAsnAspGlnLeuLys 553

```

```
Db 1347 TATGAAGAAGATTGGCAAAA-----GCTAGAGAGAGCTGAGC 1385
QY 554 GlnValGlnGlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuGlu 573
Db 1386 CGTCTACAGCAAGAAACAGACAGAAATGGAGAGAGGTGAGAGTCAGGGAAGGT----- 1439
QY 574 AlaLysGluLeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluThr 593
Db 1440 CAGTTGGAACTCTTCAGCAGCAGCCTACAAAGATTCACAAAGGAAATTAGTTCAATGCCAA 1499
QY 594 ArgSer-----LysLeuGlnGluIleAspIlePheAsnAsnGlnLeuLysGluLeuArg 611
Db 1500 ATGAACACTGATGGAATGAAGAAATTTGGAATAATCATAAATAGTCAGTTAAATTTGGTGAGT 1559
QY 612 GluIleHisAsn----- 615
Db 1560 AGCCACACAGCATTCTTGTAACGGAGCTACAGATTATTGCAGCTCAGCACCAGCAGC 1619
QY 616 -----LysGlnGlnLeuGlnLysGlnLysSerMetGluAlaGluArg 629
Db 1620 AGTGAACACAGCAACCTTAATGAACATGTTGAAGGCCAGCAGCAACCTAGAGCTCGAGCCC 1679
QY 630 LeuLysGlnLysGluGlnGluArgLysIleIleGluLeu----- 642
Db 1680 ATACACAGGAATCTCAGCAAGAGTAGTCTGGAACACTAGCTGCTTGGTGACTGAT 1739
QY 643 -----GluLysGlnLysGluGluAlaGlnArgAlaGlnGluArgAspLysGlnTrp 660
Db 1740 GAAATGAGGTGACTACAGCTGTTACTGAAAGATTGTTCTGAACTCGCAATAAT-- 1796
QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 1797 ---AGACATTCAAAGAGGAAGATCCATTAAAT-----GTAGACTCAAGT 1838
QY 681 LysLeuLysArgGluGluSerValLysLysAsp----- 692
Db 1839 TCCTCAGACGGTCAGTTGCAGATACAACTTGGAATTTTTCAGTCTGATTCCTTTGGTT 1898
QY 693 GlyGluGluLysGlyLysGlnGluAlaGlnAspLysLeuGlyArgLeu----- 708
Db 1899 GGCAGTGATCCTTTCAAG-----GATGATCCTTTTGAAAAATCGATTCATTGGT 1949
QY 709 -----PheHisGlnHisGln 713
Db 1950 GGTGATCCTTTCAAAGGTTTCAGATCCATTTGCATCAGACTGTTCTTCAGGCAATCTACT 2009
QY 714 GluProAlaLysProAlaValGlnAlaProTrpSerThrAlaGluLysGlyProLeuThr 733
Db 2010 GATCCCTTTTCCCACTTCAAGCAGCTGACCCCTTCAGTGCAGCCCAACATAGCAGTATTACA 2069
QY 734 IleSerAlaGlnGluAsnValLysValValTyrTyrArgAlaLeuTyrProPheGluSer 753
Db 2070 -----TCGGTAGAAACG 2081
QY 754 ArgSerHisaspGluIleThrIleGlnProGlyAspIleValMetValAspGluSerGln 773
Db 2082 TTGAAGCACAAATGATCCTTTTGCTCTGGTGGGAACAGTTCTTCAGCAACGATTCAGCC 2141
QY 774 ThrGlyGluProGlyTrpLeuGlyGlyGluLeuLysGlyLysThrGlyTrpPheProAla 793
Db 2142 ACAGACCCCTTGTCTGTTTTGGGAATGAATCATTTGGAGGTGATTTGCTGACTTC 2201
QY 794 AsnTyrAlaGluLysIleProGluAsnGluValPro----- 805
Db 2202 AGCACAATTGTCAAAGGTC---AACAAATGAAGATCCTTTTCGTTTCAGCCACATCGAGCTCT 2258
QY 806 -----AlaProValLysProValThrAspSerThrSerAla----- 817
Db 2259 GTCAGCAACCTAGTGTATTACAAAAAATGATTTGAGGAAACATCGGTCAAAGTGAAGAT 2318
QY 818 -----ProAla-----ProLysLeuAlaLeuArgGluThrProAlaProLeu----- 831
```

```
Db 2319 GAACCCCGAGCACTGCCACCAAGATCGGAACCTCCCAAGACCCCTGCCCTCTACCACT 2378
QY 832 -----AlaValThrSerSerGluPro----- 838
Db 2379 GGGAAAAGATCCATCAACAAATTTGGATTCTCTCTGATCCCTTTAACTGAATGATCCATTT 2438
QY 839 -----SerThrThrProAsnAsnTrpAlaAspPhe 848
Db 2439 CAGCCTTTTCCAGGCAACGATAGCCCAAGAAAAAGATCCTGAAATGTTTTGTGATCCA 2498
QY 849 SerSerThrTrpProThrSerThrAsnGluLysProGluThrAspAsnTrpAspAlaTrp 868
Db 2499 TTCATTCTGCTACTACTACCAATAAAGAGGCTGATCCAAAGCAATTTTGCACACTTC 2558
QY 869 AlaAlaGlnProSer 873
Db 2559 AGTGCTTATCCCTCT 2573

RESULT 10
US-09-404-879A-5
; Sequence 5, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Aigate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-5

Alignment Scores:
Pred. No.: 2,44e-27 Length: 531
Score: 467.00 Matches: 99
Percent Similarity: 70.66% Conservative: 19
Best Local Similarity: 59.28% Mismatches: 39
Query Match: 7.92% Indels: 10
DB: 4 Gaps: 4

US-09-720-934-2 (1-1143) x US-09-404-879A-5 (1-531)
QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 42 ATGGCTCAGTTTCCCAAGCAATGGAATGGAGGCGCAATATATGTTGGCTATTACATCTGAA 101
QY 21 GluArgAlaLysHisaspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 102 GAACTGACTAAGCATGATAAACAGTTTGATPACTCAACCTTCAGAGAGTTACATAACA 161
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 162 GGTGATCAAGCCGCTACTTTTCTTACAGTCAAGTCTCGCGCCCGCTTTAGCTGAA 221
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGCCCTTATCAGATCTGAAACAGGATGGGAAGATGGACGAGCAAGAGTTCTCTATA 281
QY 81 AlaMetLysLeuIleLysLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAACATCATCAAGTTAAAGTTGAGGCGCAACAGCTGCTGTAGTCTCCCTCCT 341
QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAACAACCCCTATGTTCTCTCCACTAATCTCTCT---CGTTTGGGATGGGA 398
```

QY 119 GlylleaSerMetProProLeuThraValaValaProValProMetGlySerIlePro 138
Db 399 -----AGCATGCCCAATCTCTCCATTCATCAGCCATTCCT-----CCA 437
QY 139 ValValGlyMetSerProThraValaValaValaProValProMetGlySerIlePro 158
Db 438 GTTGCACTATAGCAACACCTTGTCTTCTGCTACTTTCAGGACCAAGTATTCCTCCCTA 497
QY 159 AlaAsnGlyAlaProProVal 165
Db 498 ATGATGCCTGCTCCCTAGTG 518
RESULT 11
US-09-338-933-5
; Sequence 5, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338.933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-5

Alignment Scores:
Pred. No.: 2.44e-27 Length: 531
Score: 467.00 Matches: 99
Percent Similarity: 70.66% Conservative: 19
Best Local Similarity: 59.28% Mismatches: 39
Query Match: 7.92% Indels: 10
DB: 4 Gaps: 4

US-09-720-934-2 (1-1143) x US-09-338-933-5 (1-531)

QY 1 MetAlaGlnPheProThraValaValaProValProMetGlySerIleProValGlu 20
Db 42 ATGGCTCAGTTTCCACAGCATGATGAGGGGCAATATGTGGCTATTACATCTGAA 101
QY 21 GluArgAlaLysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 102 GAACGTACTAAGCATGATAACAGTTTGATAACCTCAAACTTCAGGAGGTATACATAACA 161
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValIleuAlaGln 60
Db 162 GGTATCAAGCCGCTACTTTTCTCAGTCAGTCTGCGGCCCGGTTTAGCTGAA 221
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGGCTTATCAGATCTCAAGATGGAAGATGGAGATGGACCAAGAGTTCTCTATA 281
QY 81 AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAACCTATCAGATCTCAAGATGGAAGATGGAGATGGACCAAGAGTTCTCTATA 341
QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAACCAACCCCTATGTTCTCTCCACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
QY 119 GlylleaSerMetProProLeuThraValaValaProValProMetGlySerIlePro 138
Db 399 -----AGCATGCCCAATCTCTCCATTCATCAGCCATTCGCT-----CCA 437
QY 139 ValValGlyMetSerProThraValaValaProValProMetGlySerIlePro 158
Db 438 GTTGCACTATAGCAACACCTTGTCTTCTGCTACTTTCAGGACCAAGTATTCCTCCCTA 497

QY 159 AlaAsnGlyAlaProProVal 165
Db 498 ATGATGCCTGCTCCCTAGTG 518
RESULT 12
US-09-215-681-5
; Sequence 5, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215.681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-5

Alignment Scores:
Pred. No.: 2.44e-27 Length: 531
Score: 467.00 Matches: 99
Percent Similarity: 70.66% Conservative: 19
Best Local Similarity: 59.28% Mismatches: 39
Query Match: 7.92% Indels: 10
DB: 4 Gaps: 4

US-09-720-934-2 (1-1143) x US-09-215-681-5 (1-531)
QY 1 MetAlaGlnPheProThraValaValaProValProMetGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 42 ATGGCTCAGTTTCCACAGCATGATGAGGGGCAATATGTGGCTATTACATCTGAA 101
QY 21 GluArgAlaLysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 102 GAACGTACTAAGCATGATAACAGTTTGATAACCTCAAACTTCAGGAGGTATACATAACA 161
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValIleuAlaGln 60
Db 162 GGTATCAAGCCGCTACTTTTCTCAGTCAGTCTGCGGCCCGGTTTAGCTGAA 221
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 222 ATATGGGCTTATCAGATCTCAAGATGGAAGATGGAGATGGACCAAGAGTTCTCTATA 281
QY 81 AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 282 GCTATGAACCTATCAGATCTCAAGATGGAAGATGGAGATGGACCAAGAGTTCTCTATA 341
QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 342 ATCATGAACCAACCCCTATGTTCTCTCCACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
QY 119 GlylleaSerMetProProLeuThraValaValaProValProMetGlySerIlePro 138
Db 399 -----AGCATGCCCAATCTCTCCATTCATCAGCCATTCGCT-----CCA 437
QY 139 ValValGlyMetSerProThraValaValaProValProMetGlySerIlePro 158
Db 438 GTTGCACTATAGCAACACCTTGTCTTCTGCTACTTTCAGGACCAAGTATTCCTCCCTA 497
QY 159 AlaAsnGlyAlaProProVal 165
Db 498 ATGATGCCTGCTCCCTAGTG 518
RESULT 13

US-09-216-003A-5
; Sequence 5, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; TYPE: DNA
; LENGTH: 531
; ORGANISM: Homo sapiens
US-09-216-003A-5
Alignment Scores:
Pred. No.: 2,44e-27 Length: 531
Score: 467.00 Matches: 99
Percent Similarity: 70.66% Conservative: 19
Best Local Similarity: 59.28% Mismatches: 39
Query Match: 7.92% Indels: 10
DB: 4 Gaps: 4
US-09-720-934-2 (1-1143) x US-09-216-003A-5 (1-531)
QY 1 MetAlaGlnPheProThrPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
DB 42 ATGGCTCAGTTTCCACGCGATGAATGGAGGCGCAATATATGGGCTATTACATCTGAA 101
QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
DB 102 GAACGCTACTAGCATGATACACAGTTTGTATACCTCAACCTTCAGGAGGTACATAACA 161
QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
DB 162 GGTGATCAAGCCCGTACTTTTCTTACAGTCAGGCTGCGGCGCGGTTTGTAGCTGAA 221
QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
DB 222 ATATGGGCTTATCAGATCTGAAACAGGATGGGAAGATGGACCAAGAGTTCTCTATA 281
QY 81 AlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
DB 282 GCTATGAACCTATCAAGTTAAAGTTGCAGGCGCAACAGCTGCTGTAGTCTCCCTCCT 341
QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
DB 342 ATCATGAACAACCCCTATGTTCTCTCCACTAATCTCTGCT---CGTTTGGGATGGGA 398
QY 119 GlyIleAlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIlePro 138
DB 399 -----ASCATGCCAATCTGTCTCCATTCATCAGCCATGCTCT-----CCA 437
QY 139 ValValGlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeu 158
DB 438 GTTGCACCTATAGCAACACCCCTGTCTCTGCTACTTCAGGGACAGTATCTCCCTCA 497
QY 159 AlaAsnGlyAlaProProVal 165
DB 498 ATGATGCTGCTCCCTCCTAGTG 518
RESULT 14
US-08-095-737-3
; Sequence 3, Application US/08095737
; Patent No. 5487979
; GENERAL INFORMATION:
; APPLICANT: DiFiore, Pier P
; APPLICANT: Fazioli, Francesca
; TITLE OF INVENTION: A Substrate for the Epidermal Growth

; TITLE OF INVENTION: Factor Receptor Kinase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: United States of America
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,737
; FILING DATE: 19930722
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH060.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 111..2802
; US-08-095-737-3
Alignment Scores:
Pred. No.: 5.16e-26 Length: 3033
Score: 465.00 Matches: 193
Percent Similarity: 36.01% Conservative: 152
Best Local Similarity: 20.15% Mismatches: 287
Query Match: 7.89% Indels: 326
DB: 1 Gaps: 32
US-09-720-934-2 (1-1143) x US-08-095-737-3 (1-3033)
QY 15 TrpAlaIleThrValGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 34
DB 474 TGGGCTGTAAGTCTGAAGATAAGGCAATATATGATGCAATTTTTCACAGTTTAAAGCCCA 533
QY 35 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 54
DB 534 GTGGATGAGATTTTGTCTGTGTGATAAGTGAACCAAGTGTGTCTCACTTAAGTTACCT 593
QY 55 GlnProValLeuAlaGlnIleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 74
DB 594 GTGGAAATCCTTGAAGAGTTTGGGAGTTTGGAGTGATATTGACCACGATGGAAGCTGAC 653
QY 75 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 94
DB 654 AGAGATGAGTTTGCAGTTGCCATGTTTGTGTATCTGTGCTACTGTGCTGAGGAG----- 701
QY 95 ProSerAlaLeuProProValMetLysGlnGlnProValAlaIleSerSerAlaProAla 114
DB 701 ----- 701
QY 115 PheGlyMetGlyGlyIleAlaSerMetProProLeuThrAlaValAlaProValProMet 134
DB 702 -----AAAGAACCTGTGCCAATG 719

QY 135 GlySerIleProValValGlyMetSerProThrLeuValSerSerValProThrAlaAla 154
Db 720 -----TCTTGCCTCCAGCCTTG 737
QY 155 ValProProLeuAlaAsnGlyAlaProProValIleGlnProLeuProAlaPheAlaHis 174
Db 738 GTGCCACCT----- 746
QY 175 ProAlaAlaThrLeuProLysSerSerSerPheSerArgSerGlyProGlySerGlnLeu 194
Db 747 -----TCTAAGAGA 755
QY 195 AsnThrLysLeuGlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGlu 214
Db 756 AAAACG----- 761
QY 215 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 234
Db 762 TGGGTTGTATCCCTGCAGAAAAGCTAAATATGATGAAATTTTCTGAAAACCTGATAAG 821
QY 235 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 254
Db 822 GATATGATGATGATGTCGTGGACTCGAGGTCCGTAACCTTCCTGAAAACAGGTTTA 881
QY 255 ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeu 274
Db 882 CTTTCTGCCTTGTAGCCACATTTGTCTACTATGTGCACAAAGGCTGTGGGAAGCTT 941
QY 275 ThrAlaGluPheIleLeuAlaMetHisLeuIleAspValAlaMet---SerGlyGln 293
Db 942 TCAAAAGACCACTTGGCTTGGCTTTTCACTTAATCAATCAGAAAGTTAATAAAAGGCATT 1001
QY 294 ProLeuProProValLeuProProGluTyrIleProProSerPheArgValArgSer 313
Db 1002 GACCTCTCATAGTCTACCTCCGAGATGATCCACCATCAGACAGA----- 1049
QY 314 GlySerGlyIleSerValIleSerSerThrSerValAspGlnArgLeuProGluPro 333
Db 1049 ----- 1049
QY 334 ValLeuGluAspGluGlnGlnGlnLeuLysLysIleProValThr---PheGluAsp 352
Db 1050 -----TCCAGTTTACAAAAGAACATCACAGGATCAAGTCTGTGTGCAGATTTTCTGCT 1103
QY 353 LysLysArgGluAsnPheGluArgGlyAsnLeuGluLeuLysArgArgGlnAlaLeu 372
Db 1104 ATTAAG-----GACTAGATACCTTTAACATGAATA 1136
QY 373 LeuGluGlnArgLysGluGlnGluArgLeuAlaGlnLeuGluArgAlaGluGlnGlu 392
Db 1137 GTTGACCTGCAGAGGAAAAG-----AACATGTGAGCAGGAC 1175
QY 393 ArgLysGluArgGluArgGlnGlnGluArgLysArgGlnLeuGluLeuLysGln 412
Db 1176 CTTAAAGAGAGGAGACACAGTTAAGCAGAGGACCGATGAGGTT----- 1220
QY 413 LeuGluLysGlnArgGluLeuGluArgGlnArgGluGluArgGlyGluIleGlu 432
Db 1221 -----CAGGATCTCAAGATGAATGAA 1244
QY 433 ArgArgGluAlaAlaLysArgGluLeuGluArgGlnArgGlnLeuGluIleValLeuLys 452
Db 1245 AGGAGAGATTAATCTCAAAAACCTGCAGGCCAGAAAGCAG----- 1286
QY 453 ArgArgGlnLeuLeuAsnGlnArgAsnLysGluGlnGluAspIleValValLeuLys 472
Db 1287 CAGGTGCAGGAGCTCCTGGGTGAA----- 1310
QY 473 AlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeu 492
Db 1311 -----CTGGATGAGCAGAAAGCCAGCTG 1334
QY 493 GluGlyLysLeuGlnAspIleArgCysArgLeuThrThrGlnArgGlnGluLeuGluSer 512

Db 1335 GAGGAGCAGCTCCAGGAAGTCAGAAAAGTGTCTGAGGAGGCCCCAGCTGATTCTTCTCC 1394
QY 513 ThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnGlnGln 532
Db 1395 CTGAAA-----GCAGAAATAACTAGTCAAGAACTCTCAGATCTCC 1433
QY 533 GluSerGlnGlnMetLeuGlyArgLeuIleProGluLysGlnIleLeuAsnAspGlnLeu 552
Db 1434 AGTTATGAGGAAGACTGTTGAAA-----GCTAGAGAAGAACTA 1472
QY 553 LysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeu 572
Db 1473 AGTCGCTCAACAAGAAACAGCACAATTGGAAGAAAGTGTGGAGTCAGGGAAGCT--- 1529
QY 573 GluAlaLysGluLeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLys--- 591
Db 1530 ---CAGCTGAAACCTCTTTCAGCAGCACCTACAGAGTCACAAACAGGAAATCAGCTCAATG 1586
QY 592 GluThrArgSerLysLeuGlnGluIleAspIlePheAsnAsnGln----- 606
Db 1587 CAATGAGATTGGAATGAAAGATCTGGAACCTGATAATACCAATCAAATTCGAGCAGT 1646
QY 607 -----LeuLysGlu 609
Db 1647 AGCCCAACAAGCGTTCTTGTATATGTTGTACAGATTACTGTAGCCTCAGCAGCAGCAGC 1706
QY 610 LeuArgGluIleHisAsnLysGlnGlnLeuGlnLysGlnLysSerMetGluAlaGluArg 629
Db 1707 AGTGAACAGCCAACTTCAACGAACATGCTGAAGCCCAAAACAACAGTAGCTCTGAACCC 1766
QY 630 LeuLysGlnLysGluGlnGluArgLysIleIleGluLeu-----GluLys 644
Db 1767 ACACACAGAGCTCTCAGTAAGAGTAGTCTGAAATCGCACCTTCTGATGACTGAT 1826
QY 645 GlnLysGluGluAlaGlnArgArgAlaGlnGluArgAspLysGlnTrpLeuGluHisVal 664
Db 1827 GAAAGTCAGGCTGTGACTGTGGCTGTGTAATGAGAAAGTTACTCCGAGATTGAC----- 1880
QY 665 GlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGluLysLeuLysArg 684
Db 1881 -----GATCACAAGCAGCTCAAAAGAGGAGATCCATTAATAGTAGAATCAAGTTCACTG 1934
QY 685 GluGluSerValLysLysLysAsp-----GlyGlu 694
Db 1935 ACAGATCAGTTGCAGATACAAACTTGTGATTTTCCAGTCTGATCCTTTTGTGGCAGT 1994
QY 695 GluLysGlyLysGlnGluAlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGlu 714
Db 1995 GATCCTTTCAAG-----GATGATCCTTTTGGAAAAATGATCCATTTGGTGGTGAC 2045
QY 715 Pro-----AlaLysProAlaValGlnAla 722
Db 2046 CTTTCAAGAGCTCAGATCCTTTTGGCTGTGATGCTTTTAAAGCAGACTTCTACTGAT 2105
QY 723 ProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnValLysVal 742
Db 2106 CTTTACCAGCTTCAAGTACGAGCCCTTTCAGTGCATCCAGCAACAGCAGTAATACATCG 2165
QY 743 ValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGln 762
Db 2166 GTA-----GAACTTGAAGCAGCAATGATGCCCATTTGCTCTCT 2201
QY 763 ProGlyAspIleValMetValAspGluSerGlnThrGlyGluPro-----GlyTrpLeu 780
Db 2202 GGTGGACAGCTGTGTGTCGAGCAGTGTATTCAGCCACAGACCTTTTGTCTCTCTTTC 2261
QY 781 GlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIlePro 800
Db 2262 GGAATGAATCATTTGGA---GATGGATTGCTGACTTCAGCACATATCAAGAGTCAAC 2318
QY 801 GluAsnGluValProAlaPro-----ValLys 809

Db 2319 AATGAAGATGCTTTTAACTCCTACCATATCAAGTTCTTACCAGCAGTGTGACCAATTGCAAAA 2378
Qy 810 ProValThrAspSerThrAla-----ProAla-----ProLys 821
Db 2379 CCTATTTAGAGAAACAGCCACAGAGTGAAGATGTGCTCCAGCAGTCCGCCCCAAA 2438
Qy 822 LeuAlaLeuArgGluThrProAlaPro-----LeuAlaVal 833
Db 2439 GTTGGCACTCCAAAGACCTTGGCCGCCACCCCTGGGAAAAGACCCATCAACAATTG 2498
Qy 834 ThrSerSerGluPro----- 838
Db 2499 GATCTCTTGATCCCTTAAACTGAATGATPCCATTTTCCAGCCTTTCCAGGCAATGATAGT 2558
Qy 839 -----SerThrThrProAsnAsnTTPAlaAsp---PheSerSerThrTTPProThrSer 855
Db 2559 CCAAAAGAAAAGATCTGATATGTTTGTGATCCATTCATCTCTCT-----ACCACT 2612
Qy 856 ThrAsnGluLysProGluThrAspAsnTTPAspAlaTTPAlaAlaGlnProSer 873
Db 2613 ACCAATAAGAGGCTGACCCCAAGCAATTTTGCTAACTTCAGTGTCTATCCCTCT 2666
RESULT 15
US-08-480-145-3
; Sequence 3, Application US/08480145
; Patent No. 5717067
; GENERAL INFORMATION:
; APPLICANT: DiPiero, Pier P
; APPLICANT: Fazioli, Francesca
; TITLE OF INVENTION: A Substrate for the Epidermal Growth
; TITLE OF INVENTION: Factor Receptor Kinase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: United States of America
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,145
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/095,737
; FILING DATE: 22-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH060.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 111..2802
; US-08-480-145-3

Alignment Scores:

Pred. No.:	5.16e-26	Length:	3033
Score:	465.00	Matches:	193
Percent Similarity:	36.01%	Conservative:	152
Best Local Similarity:	20.15%	Mismatches:	287
Query Match:	7.89%	Indels:	326
DB:	1	Gaps:	32
US-09-720-934-2 (1-1143) x US-08-480-145-3 (1-3033)			
Qy	15	TpAlaIleThrValGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro	34
Db	474	TGGGCTGTTAAAGTCTGAAGATAAAGCAATATGATGCAATTTTGACAGTTTAAGCCCA	533
Qy	35	IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro	54
Db	534	GTGGATGGATTTTGTCTGTGATAAAGTGAAACCAAGTGTGTCAACTTAAGTTACCT	593
Qy	55	GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp	74
Db	594	GTGGAATCCTTGGAAAGAGTTTGGGAGTTGAGTGATATTGACCACCATGGAAGAGTGAC	653
Qy	75	GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysGlnGlyTyrGlnLeu	94
Db	654	AGAGATGAGTTTGCAGTTCCTGTTGGTATCTGTCACCTGGAG-----	701
Qy	95	ProSerAlaLeuProProValMetLysGlnGlnProValAlaIleSerSerAlaProAla	114
Db	701	-----	701
Qy	115	PheGlyMetGlyGlyIleAlaSerMetProProLeuThrAlaValAlaProValProMet	134
Db	702	-----AAAGAACCTGTGCCAATG	719
Qy	135	GlySerIleProValValGlyMetSerProThrLeuValSerValProThrAlaAla	154
Db	720	-----TCTTGCCTCCAGCCTTG	737
Qy	155	ValProProLeuAlaAsnGlyAlaProProValIleGlnProLeuProAlaPheAlaHis	174
Db	738	GTGCCACCT-----	746
Qy	175	ProAlaAlaThrLeuProLysSerSerSerPheSerArgSerGlyProGlySerGlnLeu	194
Db	747	-----TCTAAGAGA	755
Qy	195	AsnThrLysLeuGlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGlu	214
Db	756	AAAACG-----	761
Qy	215	TpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys	234
Db	762	TGGTTGTATCCCTCGCAGAAAAAGCTAAATATGATGAAATTTTCTGAAACCTGATAAG	821
Qy	235	ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu	254
Db	822	GATATGGATGGATATGTCTGGAGCTCGGAGGTCGTAACCTTCTCGAACAACAGGTTTA	881
Qy	255	ProGlnAlaGlnLeuAlaSerIleTTPAsnLysSerAspIleAspGlnAspGlyLysLeu	274
Db	882	CCTTCTGCTTGTAGCCCACTTTGGTCACTATGTGACACAAAGGGCTGTGGAGACTT	941
Qy	275	ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMet---SerGlyGln	293
Db	942	TCAAAAGACCAGTTTGCCTTGGCTTTTCTTAATCAATCAGAAGTTAATAAAGGCATT	1001
Qy	294	ProLeuProProValLeuProProGluTyrIleProProPheArgValArgSer	313
Db	1002	GACCTCTCATAGTCTCACTCTGAGATGATTCCACCATCAGACAGA-----	1049
Qy	314	GlySerGlyIleSerValIleSerSerThrSerValAspGlnArgLeuProGluGluPro	333
Db	1049	-----	1049


```
QY 334 ValLeuGluAspGluGlnGlnGlnLeuGluLysLysLeuProValThr---PheGluAsp 352
Db 1050 -----TCCAGTTTACAAAGAACATCACAGGATCAAGTCTCTGTTGCAGATTTTCTGCT 1103
QY 353 LysLysArgGluAsnPheGluArgGlyAsnLeuGluLeuGluLysArgArgGlnAlaLeu 372
Db 1104 ATTAAG-----GAACTAGATACCCTTAAACAATGAAATA 1136
QY 373 LeuGluGlnGlnArgLysGluGlnGlnGluArgLeuAlaGlnLeuGluArgAlaGluGlnGlu 392
Db 1137 GTTCACTGCAGAGGAAAG-----AACATGTGCAGCAGAC 1175
QY 393 ArgLysGluArgGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 412
Db 1176 CTTAAAGAGAAGAACACAGTTAAGCAGAGGACCAGTGAGTT----- 1220
QY 413 LeuGluLysGlnArgGluLeuGluArgGlnArgGluGluGluArgArgLysGluLeuGlu 432
Db 1221 -----CAGGATCTTCAAGATGAAGTTCAA 1244
QY 433 ArgArgGluAlaAlaLysArgGluLeuGluArgGlnArgGlnArgGlnArgGlnArgGln 452
Db 1245 AGGAGAGTATTAACTACAAAACCTGCAGCCCCAGAGCAG----- 1286
QY 453 ArgArgGlnGluLeuLeuAsnGlnArgAsnLysGluGlnGlnGluAspIleValValLeuLys 472
Db 1287 CAGTGCAGGAGCTCTGGGTGA----- 1310
QY 473 AlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeu 492
Db 1311 -----CTGGATGAGCAGAAAGCCAGCTG 1334
QY 493 GluGlyLysLeuGlnAspIleArgCysArgLeuThrThrGlnArgGlnGluLeuGluSer 512
Db 1335 GAGGAGCAGCTCCAGGAGCTCAGAAAAGTGCTGAGGAGGCCAGCTGATTTCTTCC 1384
QY 513 ThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnGlnLeuGln 532
Db 1395 CTGAAA-----GCAGAAATACTAGTCAAGATCTCAGATCTCC 1433
QY 533 GluSerGlnGlnMetLeuLysArgLysLeuProGluLysGlnIleLeuAsnAspGlnLeu 552
Db 1434 AGTTATGAGGAAGAACTGTTGAAA-----GCTAGAGAAGAACTA 1472
QY 553 LysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeu 572
Db 1473 AGTCCCTTACCAACAGAAACAGCAATGGAAGTGGAGTGGAGTCAAGGAGCT--- 1529
QY 573 GluAlaLysGluLeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLys--- 591
Db 1530 ---CAGCTGGAACCTCTTACAGCAGCACCCTACAGAGTCACAACAGGAATCAGTCAATG 1586
QY 592 GluThrArgSerLysLeuGlnGluLeuAspIlePheAsnAsnGln----- 606
Db 1587 CAATATGAGATTGGAATGAAAGATCTGGAATACTGATAATACCAATCAAAATGGAGCAGT 1646
QY 607 -----LeuLysGlu 609
Db 1647 AGCCCAACAAAGCTTCTGTTAATGGTCTACAGATTACTAGCTCAGCACCAGCAGC 1706
QY 610 LeuArgGluIleHisAsnLysGlnGlnLeuGlnLysGlnLysSerMetGluAlaGluArg 629
Db 1707 AGTGAACACAGCAACTTCAACAGACATGCTGAAGGCCCAACCAACCTAGAGTCTGAACCC 1766
QY 630 LeuLysGlnLysGluGlnGluArgLysIleIleGluLeu-----GluLys 644
Db 1767 ACACACAGGAGTCTCTAGTAAGAGTAGTCTGAAATCGCACCTTCTGTGTGACTGAT 1826
QY 645 GlnLysGluGluAlaGlnArgAlaGlnGlnArgAspLysGlnTrpLeuGluHisVal 664
Db 1827 GAAAGTGAGGCTGTGCTGCTGCTGTAATGAGAAAGTTACTCCGAGATTTGAC----- 1880
QY 665 GlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluLysLysArg 684
```

```
Db 1881 -----GATCAGCAGCAGCTCAAAAAGAGAGATCCATTTAATGTAGATCAAGTTCACTG 1934
QY 685 GluGluSerValLysLysLysAsp-----GlyGlu 694
Db 1935 ACAGATCGATTCGAGATACAACTTCGATTTTTCAGTCTGATCTCTTTGTTGGCAGT 1994
QY 695 GluLysGlyLysGlnGluAlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGlu 714
Db 1995 GATCCTTTCAAG-----GATGATCCTTTTGAAAAAATGATCCATTTGGTGGTGAC 2045
QY 715 Pro-----AlaLysProAlaValGlnAla 722
Db 2046 CTTTCAAAGGCTCAGATCCTTTTTCGCTGCTGATTCCTTTTAAAGCAGACTTCTACTGAT 2105
QY 723 ProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnValLysVal 742
Db 2106 CTTTTTACCCTTCAAGTACGAGCCCTTTCAGTGCATCCCAACAGCAGCAGTAATACATCG 2165
QY 743 ValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGln 762
Db 2166 GTA-----GAAATTTGGAAGCACAATGACCCATTTGCTCCT 2201
QY 763 ProGlyAspIleValMetValAspGluSerGlnThrGlyGluPro-----GlyTrpLeu 780
Db 2202 GGTGGAACAGTTGTTGCTGCAGCAGAGTGAATTCAGCCACAGACCCTTTGCTCTGTTTTC 2261
QY 781 GlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIlePro 800
Db 2262 GGAATGAATCATTTGGA-----GATGGATTTGCTGACTTCAGCACATATCAAGGTCAAC 2318
QY 801 GluAsnGluValProAlaPro-----ValLys 809
Db 2319 AATGAAGATGCTTTTAAATCCTTACCATATCAAGTTCTACCAGCAGTGTGACCATTGCAAAA 2378
QY 810 ProValThrAspSerThrSerAla-----ProAla-----ProLys 821
Db 2379 CCTATGTTAGAGAAACAGCCAGCAGAGTGAAGATGTGCTCCAGCAGCTCCGCCCAAA 2438
QY 822 LeuAlaLeuArgGluThrProAlaPro-----LeuAlaVal 833
Db 2439 GTTGGCTCCCAACAGACCTTGCCGCCCCCTCGGAAAAGACCCCATCAACAAATG 2498
QY 834 ThrSerSerGluPro----- 838
Db 2499 GATTTCTTGATCCCTTAAACTGAATGATCCATTTTCAGCCTTTCCAGGCAATGATAGT 2558
QY 839 -----SerThrThrProAsnAsnTrpAlaAsp---PheSerSerThrTrpProThrSer 855
Db 2559 CCCAAAGAAAAGATCCTGATATGTTTGTGATCCATTCATCTTCT---ACCACT 2612
QY 856 ThrAsnGluLysProGluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSer 873
Db 2613 ACCAATAAGAGGCTGACCCCAAGCAATTTTGTCTAACTTCAGTGTCTATCCCTCT 2666
```

Search completed: August 3, 2004, 08:55:16
Job time : 268 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 3, 2004, 03:11:34 ; Search time 5689 Seconds
(without alignments)
5999.731 Million cell updates/sec

Title: US-09-720-934-2
Perfect score: 5895
Sequence: 1 MAQPTTFFGSLDIWAITVE.....QVGLFSPNVKVLTTDMDFSPQ 1143

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09720934/runat_29072004.164338.1230/app_query.fasta.1.1287
-DB=EST -QFMT=fastap -SUFFIX=Jul29.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=PCO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934 @CGN 1 1 5180 @runat_29072004.164338.1230 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrt.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5422.5	92.0	5385	11	BC062938	BC062938 Mus muscu
2	3250	55.1	2079	11	BC013578	BC013578 Homo sapi
3	3246	55.1	2126	11	BC020269	BC020269 Homo sapi
4	1859.5	31.5	2641	11	AK029325	AK029325 Mus muscu
5	1645.5	27.9	2352	11	AK034367	AK034367 Mus muscu
6	1428.5	24.2	979	13	BQ719508	BQ719508 AGENCOURT
7	1283.5	21.8	873	14	CF729634	CF729634 UI-M-HD0-
8	1269.5	21.5	911	13	BQ942708	BQ942708 AGENCOURT
9	1262.5	21.4	818	14	CA324450	CA324450 UI-M-FY0-
10	1257	21.3	886	13	BQ319490	BQ319490 603485809
11	1242	21.1	773	12	BM681943	BM681943 UI-E-E01-
12	1237.5	21.0	886	13	BQ317489	BQ317489 603849519
13	1237	21.0	968	13	BU441953	BU441953 603208112
14	1221	20.7	819	14	CF741757	CF741757 UI-M-HB0-
15	1217	20.6	751	14	CA448081	CA448081 UI-H-ED1-
16	1213	20.6	952	13	BU127625	BU127625 603114996
17	1197	20.3	776	13	EX758765	EX758765 BX758765
18	1194	20.3	907	13	BU128229	BU128229 603114019
19	1176	19.9	842	9	AL870708	AL870708 AL870708
20	1175.5	19.9	786	13	BU119880	BU119880 603142511
21	1175.5	19.9	1550	11	AK042449	AK042449 Mus muscu
22	1170	19.8	722	14	CF533008	CF533008 UI-M-FY0-
23	1168.5	19.8	732	14	CA750495	CA750495 UI-M-FY0-
24	1165	19.8	738	12	BM944544	BM944544 UI-M-EH0P
25	1162	19.7	785	13	BU750771	BU750771 CH3#035 G
26	1158	19.6	765	14	CF539131	CF539131 UI-M-GI0-
27	1149	19.5	734	9	AI671143	AI671143 wbl3g09.x
28	1143	19.4	729	13	BU704308	BU704308 UI-M-F00-
29	1134.5	19.2	853	13	EX695683	EX695683 BX695683
30	1129	19.2	736	14	CF538122	CF538122 UI-M-GI0-
31	1123	19.1	690	13	BQ179493	BQ179493 UI-M-EW0-
32	1122	19.0	896	13	BQ941336	BQ941336 AGENCOURT
33	1101	18.7	962	12	BG249478	BG249478 602319108
34	1097	18.6	729	14	CF742748	CF742748 UI-M-HB0-
35	1089.5	18.5	941	10	CB206102	CB206102 AGENCOURT
36	1089	18.5	650	10	BB656585	BB656585 BB656585
37	1082.5	18.4	706	14	CB521237	CB521237 UI-M-GH0-
38	1078.5	18.3	688	14	CB244555	CB244555 UI-M-FY0-
39	1077.5	18.3	725	14	CB248849	CB248849 UI-M-EX0-
40	1071.5	18.2	952	13	BX410112	BX410112 BX410112
41	1068.5	18.1	698	14	CF737796	CF737796 UI-M-HD0-
42	1066.5	18.1	1015	13	BQ940346	BQ940346 AGENCOURT
43	1061	18.0	687	14	CB521525	CB521525 UI-M-GH0-
44	1055	17.9	641	12	BM728997	BM728997 UI-E-E01-
45	1054	17.9	774	13	BU316793	BU316793 603855209

ALIGNMENTS

RESULT 1
BC062938
LOCUS
DEFINITION
BC062938 Mus musculus intersectin (SH3 domain protein 1A), mRNA (cdna clone IMAGE:6839463), containing frame-shift errors.
ACCESSION
BC062938
VERSION
BC062938.1
KEYWORDS
HTC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 5385)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, I., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

PUBMED

2 (bases 1 to 5385)

REFERENCE

Strausberg, R.

AUTHORS

Direct Submission

TITLE

Submitted (26-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>

Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu

Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A.,

Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,

Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J.,

Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: Plate: Row: Column: 0

This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

source

1..5385

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6839463"

/tissue_type="Brain"

/clone_lib="NIH_BMAP_GHO"

/lab_host="DH10B"

/note="Vector: pYX-ASC"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 5385

Score: 5422.50 Matches: 1060

Percent Similarity: 90.36% Conservative: 37

Best Local Similarity: 87.31% Mismatches: 44

Query Match: 91.98% Indels: 74

DB: 11 Gaps: 3

US-09-720-934-2 (1-1143) x BC062938 (1-5385)

QY 1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20

Db 262 ATGGCTCAGTTTCCACACACCTTTTCGGTGGTATGCTCGGCGCCATAAATGTGGAG 321

QY 21 GluArgAlaIysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40

Db 322 GAAAGGCCCAAGCATGACGACGATTCCTTAGCCCTGAAGCGATAGCGGGATTTATTACT 381

QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60

Db 382 GGTGATCAAGCGAGAACTTTTTTTCATCTGGTGTACCTCAGCCTGTCTTAGACAA 441

QY 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80

Db 442 ATATGGCGCTAGCGGACATGATTAACGATGAAGGATGATCAAGTGAATTTTCCATA 501

QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuPro 100

Db 502 GCCATGAAGCTTATCAAACTGAAGTCAAGGATATCAGCTCCCTCCACATTCCTCCCT 561

QY 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120

Db 562 GTCATGAACAGCAACACAGTGGCTATTTCCAGTCACACAGCATTTGGTAGAGGGATT 621

QY 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140

Db 622 GCTAGCATGCCACCTCACAGCTGTGTCTCTGCAATGGCTCCATTCACAGTTGTT 681

QY 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160

Db 682 GGAATGCTCCACCTTAGTATCTCTGCTCCCTCAGCAGCAGTCCCTCCCTGCTTAAC 741

QY 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180

Db 742 GGGGCTCTCCCGTCATACAGCTCTGCTGGCTTGGGCTCTCGGACCATTCGCA 801

QY 181 LysSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200

Db 802 AAGAGTTCTCTCTTTCAGCAGATCTGTCAGGGGTCACAATTAACACTAAGTTACAGAAG 861

QY 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220

Db 862 GCACAACTATTCGATGTCGCGAGCGCCCTCCAGCAGCAGATGGGTGCTGCTCAGTCA 921

QY 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240

Db 922 TCAAGGCTGAATATACAGCGAGTATTTCACAGCCACGACCAAACTATGAGTGACACTTA 981

QY 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260

Db 982 ACAGGTCCCGAGGCAAGAACTATTCTCATGCAATCAAGTGTACCAAGTGTGCTGCTGCT 1041

QY 261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280

Db 1042 TCAATATGGAATCTTCTGCAATTGATCAAGATGGAANAATCTCATGAGAGAAATTTATC 1101

QY 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300

Db 1102 CTAGCTATGCACCTAATTGATGTTGCCATGCTGTGTCAGCCACTGCGCGCGCTCTCGCT 1161

QY 301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320

Db 1162 CCAGAAATACATCCCTCTCTCTTTCAGAAAGATTGCTCCGCGCATGGGATGTCTCGTCATA 1221

QY 321 SerSerThrSerValAspGlnArgLeuProGluProValLeuLeuAspGluGlnGln 340

Db 1222 AGCTCTCTCTGTTGGATCAGAGGCTGCTGAGAGGCGCTGCTCAGAGGATGAGCAGCAG 1281

QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360

Db 1282 CCA---GAGAAGAAACTGCTGTGACATTTGAAGATAAAGCGGGAGAACTTCGAGCGCA 1338

QY 361 GlyAsnLeuGluLeuGluLysArgGlnAlaLeuLeuGlnGlnArgLysGluGln 380

Db 1339 GGCAGTGTGAGCTGGAGAAAGCGCGCCAAAGCGCTCTTTGAGCAGCAGCGCAAGAGCAG 1398

QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu 400
Db 1399 GAGCGGTTGGCTCAGCTGGAGCGCGCCGAGCAGGAGAGGAAAGAGCGGAGCGCCAGGAG 1458
QY 401 GlnGluArgLysArgGlnLeuGlnLeuGluLysGlnLeuGluLysGlnArgGlnLeuGlu 420
Db 1459 CAGGAGCGCAAGCGGAGCTGGAGCTGGAGAGACAGCTGGAGAGCGGAGCTGGAG 1518
QY 421 ArgGlnArgGluGluArgLysGluLeuGluArgGlnAlaAlaLysArgGlu 440
Db 1519 CGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1578
QY 441 LeuGluArgGlnArgGlnLeuGluArgGlnArgGlnArgGlnLeuLeuLeuGln 460
Db 1579 CTGGAAGGCGGAGCAACTTGAATGGGAACGGAAACGGGAGACAGGAACTCTGTAATCAG 1638
QY 461 ArgAsnLysGluGlnGluAspLysValValLeuLysAlaLysLysLysThrLeuGluPhe 480
Db 1639 AGGAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1698
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGluLysLeuGlnAspLysArg 500
Db 1699 GAGTTAGAGGCTCTGTAATGCAAAAGCAATCAGTAGAAGGAAACTTCAGGATATCAGG 1758
QY 501 CysArgLeuThrThrGlnArgGlnGluLeuGluSerThrAsnLysSerArgGluLeuArg 520
Db 1759 TGTGACTGGCAACCCAGAGGCAAGAAATTGAGAGCACCAAGCTTAGAGAGCTAAGA 1818
QY 521 IleAlaGluLeuThrHisLeuGlnGlnGlnLeuGlnGlnSerGlnMetLeuGluArg 540
Db 1819 ATTGCTGAATCACCACCTTACAGCAGCAGTGTGAGGAACTCTCAGAAATGCTTGAAGA 1878
QY 541 LeuLeuProGluLysGlnLeuLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1879 CTTATTCCAGGAAACAGATACCTCAGTGACCACTTAAACAGTCCAGCAGAACAGTTTG 1938
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1939 CATAGAGACTCGCTTCTTACCCTCAAAAGAGCCTTGGAAAGAGAGCTGGCCGCGCAG 1998
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluLe 600
Db 1999 CAGCTCCGGAGCAGCTGGAGCAGGCTGGAGAGAGACCAAGCTCAAGCTGCAGGAGATT 2058
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluLeuHisAsnLysGlnGlnLeuGln 620
Db 2059 GATGTTTTCAACCAACCACTGAAGGAACTGAGAGAGATACATAGCAAAACAGCAACTCCAG 2118
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnArgLysIleIle 640
Db 2119 AAGCAGAGGTCCTGGAGGCGCGCAGCTGAAGCAGAAAGAGCAGGAGAGAGAGCGCTG 2178
QY 641 GluLeuLysGlnLysGluGluAlaGlnArgAlaGlnGluArgAspLysGlnTrp 660
Db 2179 GAGTTAGAGAAAGGAAAGAGCTGTCAGAGACGAGTTTCAGGAAAGGGAACAAGCAATGG 2238
QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
Db 2239 CTGGAGATGTGCACGAGGAG---GAGCAGCCAGCCCCCGGAAACCCACAGGAGGAGC 2295
QY 681 LysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLysGlyLysGlnGlu 700
Db 2296 AGACTCAAGAGGGAAGACAGTGTGAGGAAGAGGAGGCGGAGAGAGAGCAGGAGCAGGAA 2355
QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGlnProAlaLysProAlaVal 720
Db 2356 ATCAAGACAGCAGAGCTGCGCTTTTCCATCCGATCAGGAGCGAGCTAAGCTGGCCACC 2415
QY 721 GluAlaProTrpSerThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
Db 2416 CAGCACCCCTGGTCTACACAGAGAAAGGCCCTT-ACCATTCTGCACAGGAGAGTGTA 2474

QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
Db 2475 AAAAGTGGTATATTACCGAGCGCTGTACCCCTTTGAATCAGAAAGTCACGATGAGATCAC 2534
QY 761 IleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeu 780
Db 2535 ATCCAGCAGGAGATATAGTCATGTTGGATGAAGCCAGACTGGAGAGCCAGATGGCTT 2594
QY 781 GlyGlyGluLeuLysGlyLysThrGlyTrpPheProAlaAsnTyrAlaGluLysIlePro 800
Db 2595 GGAGCAGAGCTCAAAAGGGAAGACGGATGGTTCCCTGCAAACTATGCAAGAAAGATTCCA 2654
QY 801 GluAsnGluValProAlaProValLysProValThrAspSerThrSerAlaProAlaPro 820
Db 2655 GAAATATGAGGTTCCCACTCCAGCCAAACAGTGACCGATCTGCACATCTGCCCTTCCCCC 2714
QY 821 LysLeuAlaLeuArgGluThrProAlaProLeuAlaValThrSerSerGluProSerThr 840
Db 2715 AAACCTGGCTCTCGGTGAGACCCCTGCTCTTTGCCAGTACCTCTTCTGAGCCCTCCACA 2774
QY 841 ThrProAsnAsnTrpAlaAspPheSerSerThrTrpProThrSerThrAsnGluLysPro 860
Db 2775 ACCCCCAACAACCTGGGAGACTTCAGTTCACGTGGCCCGCAGCAGCTCAAAACGAGAAAGCA 2834
QY 861 GluThrAspAsnTrpAspAlaTrpAlaAlaGlnProSerLeuThrValProSerAlaGly 880
Db 2835 GAAACGGCAACCTGGGATACATGGGCGCTCAGCCCTTCTGACCGTACTAGTGTCTGGC 2894
QY 881 GlnLeuArgGlnArgSerAlaPheThrProAlaThrAlaThrGlySerSerProSerPro 900
Db 2895 CAGTTACGCGCAGAGATCAGCCCTTTACCCAGCAGCAGCAGCAGCTGCTCTCCCATCTCCC 2954
QY 901 ValLeuGlyGlnGlyGluLysValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrpArg 920
Db 2955 GTCTGGGCGCAGGCTGAAAGGTGGAAGGGCTACAAAGCGCAAGCCCTGTATCCCTGGAGA 3014
QY 921 AlaLysLysAspAsnHisLeuAsnPheAsnLysAsnAspValIleThrValLeuGluGln 940
Db 3015 GCCAAAAAGAGCAACACCTTAAATTTTAAACAAAGTGAGCTCATCACCCTCTTGGAAACAG 3074
QY 941 GlnAspMetTrpTrpPheGlyGluValGlnGlnLysGlyTrpPheProLysSerTyr 960
Db 3075 CAAGACATGTGGTGGTTGGAGAGTTCAAGGTTCAGAGGGTGGTTCCTCCAGTCTTAC 3134
QY 961 ValLysLeuLysSerGlyProIleArgLysSerThrSerMetAspSerGlySerSerGlu 980
Db 3135 GTGAAACTCATTTACGGGCGCGTAAGGAAATCCACAGCATCATGATCTGSCCTACTGAA 3194
QY 981 SerProAlaSerLeuLysArgValAlaSerProAlaAlaLysProValValSerGlyGlu 1000
Db 3195 AGTCTGCTAGTCTAAAGAGAGTGGCTTCCCGGCCCGCAAGCCAGCCATTCCTCCGAGAA 3254
QY 1000 ----- 1000
Db 3255 GAGTTATTGCCATGTACATACAGAGATTCTGAGCAAGGAGATTAACTTTACAGCA 3314
QY 1000 ----- 1000
Db 3315 GGGGATGTGATTGTGTACTACAAAGAAAGATGGTGAAGTGGTGGACGGGAAACGGTGGCGAC 3374
QY 1000 ----- 1000
Db 3375 AAGTCCGGAGTCTTCCCTCTAACTATGTAGAGCTTAAAGATTACAGAGGGCTCTGGAAC 3434
QY 1001 ----- 1001
Db 3435 GCTGGGAAACACAGGAGTTAGGAAAAAACCTGAAATTCGCCAGGTTATTGCTCTCTAC 3494
QY 1010 ThrAlaThrGlyProGluGlnLeuThrLeuAlaProGlyGlnLeuLeuLeuLeuArgLys 1029
Db 3495 GCTGCTACTGGTCCCGAAACAACTCACCTGGCTCTCTGGGAGCTGATTCTGTATCCGAGAA 3554
QY 1030 LysAsnProGlyGlyTrpTrpGluGlyGluLeuAlaArgGlyLysLysArgGlnIle 1049

Db 3555 AAGACCAGTGGATGGTGGAGGAGAACTGCAAGCTCGAGGAAAAAGCGCCAGTA 3614
 Qy 1050 GlyTyrPheProAlaAsnTyrValLysLeuSerProGlyThrSerLysIleThrPro 1069
 Db 3615 GGGTGGTTTCAGCAAAATTATGTCAAACTTCTAAGCCCGGAAACAAGCAAAATCACCCCA 3674
 Qy 1070 ThrGluProProLysSerThrAlaLeuAlaValCysGlnValIleGlyMetTyrAsp 1089
 Db 3675 ACTGAGTACTCCCAAGACCGCAGTGCAGCCAGCAGTGTGCCAGGTGATCGGATAGCAT 3734
 Qy 1090 TyrThrAlaGlnAsnAspGluLeuAlaPheAsnLysGlyGlnIleIleAsnValLeu 1109
 Db 3735 TACACGCCCAACGATGACGACTAGCTTACGAAAGCCAGATCATCAAGTCTC 3794
 Qy 1110 AsnLysGluAspProAspTyrTrpLysGlyGluValAsnGlyGlnValGlyLeuPhePro 1129
 Db 3795 AACAAAGAGACCCGAGCTGTGGAAAGGAGAGTCAAGTGGCAAGTTGGCTCTTCCCA 3854
 Qy 1130 SerAsnTyrValLysLeuThrAspMetAspProSerGln 1143
 Db 3855 TCCAATTATGTAAGCTGACACAGACATGACCCCGCCAG 3896

RESULT 2

BC013578

LOCUS

BC013578 2079 bp mRNA linear HTC 04-MAR-2003
 DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone
 IMAGE:3878242, mRNA.

ACCESSION

BC013578

VERSION

BC013578.1

KEYWORDS

GI:15488896

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2079)

Strausberg, R.

Direct Submission

Submitted (04-SEP-2001)

National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NTH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: DCTD/DTP/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>Contact: (Dickson, Mark) mcdpaxil@stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

REMARK

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAC Plate: 14 Row: C Column: 5

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4504796

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers

1..2079

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3878242"

/tissue type="Lung, large cell carcinoma"

/clone_lib="NIH MGC_68"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:
 Pred. No.: 1..11e-199 Length: 2079
 Score: 3250.00 Matches: 646
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 99.69% Mismatches: 0
 Query Match: 55.13% Indels: 0
 DB: 11 Gaps: 0

US-09-720-934-2 (1-1143) x BC013578 (1-2079)

Qy 1 MetAlaGlnPheProThrPheGlySerLeuAspIleTPrAlaIleThrValGlu 20
 Db 132 ATGGCTCAGTTTCCAAACACCTTTTGTGGCAGCTGATATCTGGGCCATAAAGTGTAGAG 191
 Qy 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 192 GAAAGAGCGAAGCATGATCAGCAGTTCATAGTTTAAAGCCAAATATCTGGATTCATTACT 251
 Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 252 GGTGATCAAGCTAGAAACCTTTTTCATCTGGGTACCTCAACCTGTTTTCAGCAG 311
 Qy 61 IleTPrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
 Db 312 ATATGGCAGCTAGCTGACATGAATATATGATGGAAGATGATCAAGTGGAGTTTCCATA 371
 Qy 81 AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 372 GCTATGAACTTATCAAACTGAAGCTACAAAGGATATCAGCTACCTCTGCACCTTCCGCC 431
 Qy 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
 Db 432 GTCATGAAACAGCAACCGATTGCTATTTCAGGCACACAGCATTTGGTATGGAGGTATC 491
 Qy 121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
 Db 492 GCCAGCATGCCACCGCTTACAGCTGTGTCTCCAGTGCCTCAATGGGATCCATTCCAGTGT 551
 Qy 141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
 Db 552 GGAATGTCTCAACCTTAGTATCTTCTGTTCACACAGCAGCTGTGCCCCCTCCCTGGCTAAC 611
 Qy 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
 Db 612 GGGGCTCCCTCTGTATACAACTCTGCTGCTGCTCATCTCTGAGCCACACATTGCCA 671
 Qy 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
 Db 672 AAGAGTTCTTCTTTTAGTAGATCTGTCAGGGGTCAACAACCTAAACACTAAATACAAAAG 731
 Qy 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTPrAlaValProGlnSer 220
 Db 732 GCACAGTCATTTGATGTGGCCAGTGTCCACAGTGGCAGAGTGGGTGTTCCTCAGTCA 791
 Qy 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
 Db 792 TCAAGACTGAAATACAGCGCAATATTCAATAGTCATGACAAACTATGAGTGGACACTTA 851
 Qy 241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
 Db 852 ACAGGTCCCCAAGCAGAACTATTCTTATGCACTCAAGTTTACCACAGGTCTCAGCTGGCT 911
 Qy 261 SerIleTPrAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
 Db 912 TCAATATGGAATCTTTCTGCACTTGTCAAGATGGAATACTTACAGCAGAGGAATTTATC 971
 Qy 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
 Db 972 CTGGCAATGCACCTCATTTGATGTAGTATCTGTCTGGCCCAACCACTGCCACCTGCTCCTGC 1031
 Qy 301 ProGluTyrIleProProSerPheArgValArgSerGlySerGlyIleSerValIle 320
 Db 1032 CCAGATATACATCCACCTCTTTTAGAAGAGTTCGATCTGGCAGTGGTATATCTGTCTATA 1091

```

QY 321 SerSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAspGluGlnGln 340
Db 1092 AGCTCAACATCTGTAGATCAGAGCTACACAGAGAAACCACTTTTGAAGATGAACAA 1151
QY 341 GlnLeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArg 360
Db 1152 CAATTAGAAAAGAAATTTACTCTGATCTTGAAGATAAGAGCGGAGAACTTTGAACGT 1211
QY 361 GlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
Db 1212 GGCACCTGGAACTGGAGAAACAGAGCAAGCTCTCTGGAACAGACGCGCAGAGCAG 1271
QY 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu 400
Db 1272 GAGCGCTGGCCAGCTGGAGCGGGCGGAGCAGAGAGGAGGAGCGCTGAGCGCCAGGAG 1331
QY 401 GlnGluArgLysArgGlnLeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
Db 1332 CAAGAGCGCAAAAGACAACCTGGAACTGGAGAGCAATTTGAGAGCGGAGCTGCAAAACGGGA 1391
QY 421 ArgGlnArgGluGluGluArgArgLysGluLeuGluArgGluAlaAlaLysArgGlu 440
Db 1392 CGGCAGAGAGAGAGAGAGAGAGAAATTTGAGAGCGGAGGCTGCAAAACGGGA 1451
QY 441 LeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlnLeuLeuAsnGln 460
Db 1452 CTTGAAGGCAACGACAACCTGAGTGGAAACGGAATCGAAGGCAAGAACTACTAAATCAA 1511
QY 461 ArgAsnLysGluGlnGluAspLysLeuValLeuLysAlaLysLysLysThrLeuGluPhe 480
Db 1512 AGAAACAAAGAACAGAGGACATAGTTGTACTGAAAGCAAGAAAGAACTTTGGAATTT 1571
QY 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGluLysLeuGlnAspLysArg 500
Db 1572 GAATTTAGAGCTCTTAATGNTAATAAAGCATCACTAGAGGGAATCTCAAGATATCAGA 1631
QY 501 CysArgLeuThrThrGlnArgGlnGluLeuGluSerThrAsnLysSerArgGluLeuArg 520
Db 1632 TGTGATTTGACCCCAAGGCAAGAAATTTGAGAGCAACAAATCTAGAGAGTTGAGA 1691
QY 521 IleAlaGluIleThrHisLeuGlnGlnGlnGlnGlnGlnSerGlnGlnMetLeuGluArg 540
Db 1692 ATTGCCGAATCCCATCTACAGCAACAATTTACAGGAATCTCAGCAAAATGCTTTGGAAGA 1751
QY 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1752 CTTATTTCCAGAAAACAGATACTCAATGACCAATTAACAAGTTTACAGCAGACAGTTTG 1811
QY 561 HisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1812 CACAGAGATTCACCTTGTACTACTTAAAGAGCGCTTAGAAGCAAAAGAACTAGCTCGGCAG 1871
QY 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluLe 600
Db 1872 CACTCAGAGACCAACTGGATGAGTGGAGAAAGAACTAGATCAAAACTACAGAGATTT 1931
QY 601 AspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
Db 1932 GATATTTCAATTAATCAGCTGAGGAACTAAGGAATACACAATAAGCAACAACCTCCAG 1991
QY 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIle 640
Db 1992 AAGCAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAAAGAAACGAAATCATA 2051
QY 641 GluLeuGluLysGlnLysGluGlu 648
Db 2052 GAATTAGAAAACAAAAA 2075

```

RESULT 3

BC020269

LOCUS

DEFINITION

2126 bp

mRNA

linear

HTC

19-DEC-2001

mrna.

ACCESSION

BC020269

VERSION

BC020269.1

KEYWORDS

HTC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2126)

AUTHORS

Strausberg,R.

TITLE

Direct Submission

JOURNAL

Submitted (19-DEC-2001)

REMARK

NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

Schein, Duane Smalrus, Michael Smith, Lorraine Spence, Jeff Scott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAL Plate: 40 Row: n Column: 1

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4504796

This clone has the following problem: no cloning site /

microdeletion.

FEATURES

Location/Qualifiers

1..2126

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4899011"

/tissue_type="Pancreas, epithelioid carcinoma"

/clone_lib="NIH MGC 42"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

ORIGIN

Alignment Scores:

Pred. No.: 2,06e-199

Score: 3246.00

Length: 2126

Matches: 645

Percent Similarity: 100.00%

Conservative: 3

Best Local Similarity: 99.54%

Mismatches: 0

Query Match: 55.06%

Indels: 0

DB: 11

Gaps: 0

US-09-720-934-2 (1-1143) x BC020269 (1-2126)

Qy 1 MetAlaGlnPheProThrPhePheGlySerLeuAspLysTrpAlaIleThrValGlu 20

Db 182 ATGGCTCAGTTTCCAAACACCTTTTGGTGGCAGCTGATATCTGGGCCATAACTGTAGAG 241

Qy 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProLysSerGlyPheIleThr 40

Db 242 GAAAGAGCGAAGCATGATCAGCAGTTCATAGTTTAAAGCCAAATCTGGATTCAITACT 301

Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60

Db 302 GGTGATCAAGCTAGAACTTTTTCATCTGGGTACTCTCACTCACTGTTTAGCACAG 361

Qy 61 IletpAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 362 ATATGGCACTAGCTGACATGAATATGATGGAGATGGATCAAGTGGAGTTTCCATA 421
Qy 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 422 GCTATGAACCTTATCAAACTGAAGCTACAGGATATCAGCTACCCCTCTGCCCT 481
Qy 101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db 482 GTCATGAACCAACCAACCACTGCTATTCTAGCCACACAGCATTTGGTATGGAGGTATC 541
Qy 121 AlaSerMetProProLeuThraAlaValAlaProValProMetGlySerIleProValVal 140
Db 542 GCCAGCATGCCCGCTTACAGCTGTGCTCCAGTGCATGGATGCCATTCAGATTGTT 601
Qy 141 GlyMetSerProThraLeuValSerSerValProThraAlaValProProLeuAlaAsn 160
Db 602 GGAATGCTCCAACTAGTATCTTCTGTTCCACAGCAGCTGTGCCCGCCCTGGCTAAC 661
Qy 161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThraLeuPro 180
Db 662 GGGCTCCCCCTGTATACAACTCTGCTGCTGCTATGCTATCTCTGAGCCACATGGCA 721
Qy 181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThraLysLeuGlnLys 200
Db 722 AAGAGTCTCTCTTTAGTAGATCTGGTCCAGGGTCACAACTAAACATAAATTACAAAG 781
Qy 201 AlaGlnSerPheAspValAlaSerValProProValAlaGluThraAlaValProGlnSer 220
Db 782 GCACAGTCATTTGATGTGGCCAGTGTCCACAGTGGCAGAGTGGGCTGTCTCCAGTCA 841
Qy 221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThraMetSerGlyHisLeu 240
Db 842 TCAGACTGAAATACAGCAATTTCAATAGTATGATGACAAACTATAGTGGACACTTA 901
Qy 241 ThrGlyProGlnAlaArgThraIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
Db 902 ACAGGTCCCCAAGCAGAACTATCTTATGAGTCAAGTTTACCACAGGCTCAGTGGCT 961
Qy 261 SerIleThraAsnLeuSerAspIleAspGlnAspGlyLysLeuThraAlaGluPheIle 280
Db 962 TCAATATGGAATCTTCTGACATGATCAAGATGAAACTTACAGCAGAGGAAATTTATC 1021
Qy 281 LeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProProValLeuPro 300
Db 1022 CTGGCAATGCACCTCATTTGATGTAGTATGTGTCGCCAACCTGCCACCTGTCTGCT 1081
Qy 301 ProGluTyrIleProProSerPheArgArgValArgSerGlySerGlyIleSerValIle 320
Db 1082 CCAGATACATTCACCTCTTTTAGAAGAGTTCGATCTGCAGTGGTATATCTGTCTATA 1141
Qy 321 SerSerThraSerValAspGlnArgLeuProGluProValLeuGluAspGlnGln 340
Db 1142 AGCTCAACATCTGTAGATCAGAGGCTACAGAGGAACAGGTTTGTAGAGATGAACAA 1201
Qy 341 GlnLeuGluLysLeuProValThraPheGluAspLysLysArgGluAsnPheGluArg 360
Db 1202 CAATTAGAAAGAAATTTACCTGTACGTTTGAAGATAGAGAGCGGAGAACTTTGAAGCT 1261
Qy 361 GlyAsnLeuGluLysArgGlnAlaLeuLeuGluGlnGlnArgLysGluGln 380
Db 1262 GGCACTGGAACTGGAGAAAGCAAGGCAAGCTCTCTCTGGAACAGCAGCGCAAGAGCAG 1321
Qy 381 GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlu 400
Db 1322 GAGCGCTTGGCCAGCTGGAGCGGCGGAGCAGGAGAGGAGGAGCGGAGCGCGAGGAG 1381
Qy 401 GlnGluArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGluLeuGlu 420
Db 1382 CAAGAGCCCAAGACAACTGGAATCTGGAGAGCAACTGGAAAGACAGCGGAGCTAGAA 1441
Qy 421 ArgGlnArgGluGluArgGlyLysGluIleGluArgArgGluAlaAlaLysArgGlu 440

Db 1442 CGGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1501
Qy 441 LeuGluArgGlnArgGlnLeuGluTyrGluArgAsnArgArgGlnGlnLeuLeuAsnGln 460
Db 1502 CTTGAAGGCAACGACCACTTGTAGTGGGACGGAAATCGAAGGCAAGAACTACTAAATCA 1561
Qy 461 ArgAsnLysGluGlnGluAspIleValValLeuLysAlaLysLysLysThraLeuGluPhe 480
Db 1562 AGAAACAAAGAACAGAGGACATAGTTGTACTGAAGCAAGAAAGAAAGACTTTGGAATTT 1621
Qy 481 GluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArg 500
Db 1622 GAATTGAAGCTCTAAATGATAAAAGCATCACTAGAGGAACTTCAAGATATCAGA 1681
Qy 501 CysArgLeuThraGlnArgGlnGluIleGlnSerThraLysLysSerArgGluLeuArg 520
Db 1682 TGTGATTTGACCCCAAGGCAAGAAATTTGAGAGCAAAACAAATCTAGAGAGTTGAGA 1741
Qy 521 IleAlaGluIleThraHisLeuGlnGlnGlnLeuGlnSerGlnGlnMetLeuGlyArg 540
Db 1742 ATTGCCGAATCACCATCTACAGCAACAATTTACAGGAATCTCAGCAAAATCTTTGGAAGA 1801
Qy 541 LeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeu 560
Db 1802 CTTATTCCAGAAAAACAGATCTCAATGACCAATTAACCAAGTTTCAAGCAGACAGTTTG 1861
Qy 561 HisArgAspSerLeuValThraLeuLysArgAlaLeuGluAlaLysGluLeuAlaArgGln 580
Db 1862 CACAGAGATTCACTTTGTACTTCTTAAAGAGCCTTTAGAAGCAAAAGAACTAGCTCGGAG 1921
Qy 581 HisLeuArgAspGlnLeuAspGluValGluLysGluThraArgSerLysLeuGlnGluIle 600
Db 1922 CACCTACAGACCACTGGATGAAGTGAGAGAAAGAACTAGATCAAACTACAGAGATT 1981
Qy 601 AspIlePheAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGln 620
Db 1982 GATATTTCATTAATTCAGCTGAAGGAACTTAAGAGAAATACAAATAAGCAACCACTCCAG 2041
Qy 621 LysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIle 640
Db 2042 AAGCAAAAGTCCATGGAGGCTGAACGACTGAACAGAAAGAAAGCAAGAAAGATCAT 2101
Qy 641 GluLeuGluLysGlnLysGluGlu 648
Db 2102 GAATTAGAAAAAAGAAAAA 2125
RESULT 4
AK029325 2641 bp mRNA linear HTC 18-SEP-2003
LOCUS Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
DEFINITION library, clone:4832424L08 product:SH3 domain protein 1B, full
insert sequence.
ACCESSION AK029325
VERSION AK029325.1 GI:26081281
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374


```

11042159
3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
  REFERENCE Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
  AUTHORS Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
  Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
  Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matabiki,M.,
  Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
  Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.
  TITLE RIKEN integrated sequence analysis (RISA) system--384-format
  JOURNAL sequencing pipeline with 384 multipillar sequencer
  MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
  PUBLISHED 20530913
  PUBLISHED 11076861
  REFERENCE 4
  AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
  TITLE FANTOM Consortium.
  JOURNAL Functional annotation of a full-length mouse cDNA collection
  REFERENCE Nature 409, 685-690 (2001)
  AUTHORS 5
  TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research
  JOURNAL Group Phase I & II Team.
  REFERENCE Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
  JOURNAL Nature 420, 563-573 (2002)
  REFERENCE 6 (bases 1 to 2641)
  AUTHORS Adachi,S., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
  Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
  Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
  Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
  Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
  Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
  Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
  Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
  Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
  Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
  Takeda,Y., Tanaka,T., Tomaru,A., Toyota,T., Yasunishi,A.,
  Muramatsu,M., and Hayashizaki,Y.
  TITLE Direct Submission
  JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
  Physical and Chemical Research (RIKEN), Laboratory for Genome
  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
  RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
  Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
  URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
  Fax:81-45-503-9216)
  COMMENT cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
  Division of Experimental Animal Research in Riken contributed to
  prepare mouse tissues.
  Please visit our web site for further details.
  URL:http://genome.gsc.riken.go.jp/
  URL:http://fantom.gsc.riken.go.jp/.
  FEATURES
  Location/Qualifiers
  source
    1..2641
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="FANTOM:DB:4832424L08"
    /db_xref="MG:2391398"
    /db_xref="taxon:10090"
    /clone="4832424L08"
    /tissue_type="head"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stage="0 day neonate"
  misc_feature
    1..2641
    /note="SH3 domain protein 1B (MGD|MG1:1338049)"
  ORIGIN
  Alignment Scores:
  Pred. No.: 1-19e-109 Length: 2641
  Score: 1859.50 Matches: 422

```

```

Percent Similarity: 67.04% Conservative: 119
Best Local Similarity: 52.29% Mismatches: 163
Query Match: 31.54% Indels: 103
DB: 11 Gaps: 24

US-09-720-934-2 (1-1143) x AK029325 (1-2641)

Qy 1 MetAlaGlnPheProThrPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db 208 ATGGCTCAGTTTCCACAGCGATGAATGGAGGGCCAAATATGTGGCTATTACATCTGAA 267
Qy 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db 268 GAACGTACTAAGCATGATAACAGTTTGTAAAGCTCAACCTTCAGGAGGTTCATAACA 327
Qy 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db 328 GGTGATCAAGCCCGTACTTTTTTCTACGTACAGTCTGCCGCCCGCGTTTGTAGCTGAA 387
Qy 61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db 388 ATATGGCCCTTATCAGATCTGAACAGATGGAGATGGACGACCAAGAGTTCTCTATA 447
Qy 81 AlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db 448 GCTATGAACATCATCAAGTTAAAGTTTGCAGGGCCCAACAGCTGCTGTAGTCTCCTCCT 507
Qy 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
Db 508 ATCATGAACAACCCCTTAAGTTTCTCTCCACTAATCTCTGCT---CGTTTGGGATGGGA 564
Qy 119 GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaAlaProVal 132
Db 565 -----AGCATGCCAATCTGTCATTCATCAGCAATTCCTCCAGTTGCCACCTATA 615
Qy 133 -----ProMetGly-----SerIleProValValGlyMetSerPro 144
Db 616 GCACACCCCTTGCTTCTGCTACTTCAGGGACCAAGTATTCCTCCCTTAATGAGCTGCT 675
Qy 145 ThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsnGlyAlaProPro 164
Db 676 CCCTTAGTGCTTCTGTTAGTACATCTCTCAATTACCA-----AATGGAATGCCAGT 726
Qy 165 ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSerSer 184
Db 727 CTCATTGAGCCTTTA--TCCAATTCCTTATCTTCTTCAACATTCGCTCATGATCATCT 783
Qy 185 PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLysAlaGln 202
Db 784 TACAGCCTGATGATGGGAGGATTTGCT-----GGTGCTAGTATCCAGAGGCCCCAG 834
Qy 203 Ser---PheAspValAlaSerValProPro-----211
Db 835 TCTCTGATTGATTAGGATCTAGTACTCACTTCTCACTGCTTCCCTCTCCAGGGAAC 894
Qy 212 -----ValAlaGluTrpAlaValProGlnSerSerArgLeuLysIleArg 226
Db 895 TCACCTAAGACAGGGACCTCAGAGTGGGAGTTCTCTCAGCCTTCAAGATTTAAAGTATCG 954
Qy 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 955 CAAAAATTTAATAGCTAGACAAAGGCATGAGCGGATACCTCTCAGGTTTTCAGGTAGA 1014
Qy 247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 1015 AATGCCCTTCTTTCAGTCAATCTCTCTCAAACTACGTAGTACTACTATTGGACTCTGGCT 1074
Qy 267 AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle 286
Db 1075 GACATCGATGGTGACGACAGTGTGAAGCTGAAGAAATTTATTCTGGCGATGCACCTCACT 1134
Qy 287 AspValAlaMetSerGlyGlnProLeuProValLeuProGluTyrIleProPro 306
Db 1134

```

Db 1135 GACATGCCAAAGCTGGACAGCCACTACCCTGAGCTGGCTCCCGAGCTTGCTCCCTCCA 1194
 QY 307 SerPheArgValArgSerGlySerGlyIleSerValIleSerThrSerValAsp 326
 Db 1195 TCITTC-----ACAGGGGGAAGCAAGATTGATTCTGTTAATGGAAT----- 1236
 QY 327 GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnLeuGluLysLeu 346
 Db 1237 -----CTGCTTCATATCAGAAACACAGAGAGAGCTCAG-----AAGAACTG 1284
 QY 347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGlyAsnLeuGluLeuGlu 366
 Db 1285 CCAGTTACTTTTGGAGCAACAGCAAGCCAACTATCAACGAGGAAACATGAGCTGGAG 1344
 QY 367 LysArgArgGlnAlaLeuLeuGlnGlnArgLysGluGlnGluArgLeuAlaGlnLeu 386
 Db 1345 AAGCCAGCCCAAGTGTGATGAGCAGCAGCAGAGGAGGCTGAACCAAGCCCAAGAA 1404
 QY 387 GlnArgAlaGluGlnGluArgLysGluArgGlnGlnGlnGluArgLysArgGln 406
 Db 1405 GAGAGGAAGAGTGGGAGCGGAAACAGAGAGACTGCAAGAGCAAGATGGAAGAGCAG 1464
 QY 407 LeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGluArgGlnGluGlu 426
 Db 1465 CTGAGTTGGAGAAACGTTTGAGAAACAGAGAGAGCTGGAGAGACAGCGGAGAGAG 1524
 QY 427 ArgArgLysGluIleGluArgArgGluAlaLysArgGluLeuGluArgGlnArgGln 446
 Db 1525 AGGAGAAAGAGATAGAGAGCAGAGCAGCAAAACAGAGAGCTTGAGAGCAACGCGT 1584
 QY 447 LeuGluTrpGluArgAsnArgArgGlnGluLeuLeuLeuGlnArgAsnLysGluGlnGlu 466
 Db 1585 TTGAATGGGAAAGACTCCGTCGGCAGGAGCTGCTCAGTCAGAGACAGCAGGAGCAAGAA 1644
 QY 467 AspIleValIleLysAlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsn 486
 Db 1645 GACATTGTCAGCTGAGCTCCAGAAAGAAAGTCTCCACCTGGAATCGGAAGCAGTGAAT 1704
 QY 487 AspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArgCysArgLeuThrGln 506
 Db 1705 GGAAACATCATCAGCATCTCAGGCAGACTACAAGATGTCCAAATCAGAAAGCAACACAA 1764
 QY 507 ArgGlnGluIleGluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHis 526
 Db 1765 AAGATGAGCTAGAGATTTTGGATAACAGCTGACCTGGCAATTTATGGAATCAACAA 1824
 QY 527 LeuGlnGlnGlnLeuGlnGlnSerGlnGlnMetLeuGlyArgLeuIleProGluLysGln 546
 Db 1825 CTTCAACAAGAGCTTAAGCAATATCAAAATAAGCTTATCTATCTGCTCCCTCAGAGCAG 1884
 QY 547 IleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuVal 566
 Db 1885 CTATTAAACGAAAGAAATTAACATGACGCTCAGTAACACA---CCTGATTTCAGGATC 1941
 QY 567 ThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAsp 584
 Db 1942 AGTTTACTTCAATAAAGCTCATCAAGAAAGGAAGATTAATGC---CAAAGCTTAAGAA 1998
 QY 585 GlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluIleAspIlePheAsn 604
 Db 1999 CAATTAGATGCTCTTGAAGAAAGAACTGCATCTAAGCTCTCAGAAATGATTCAITTAAC 2058
 QY 605 AsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGlnLeuGlnLysSer 624
 Db 2059 AATCAGCTGAAGGAACCTCAGAAAGCTTATAATACAGCAGTTAGCCCTTGCAACACTT 2118
 QY 625 MetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLysIleIleGlu----LeuGlu 643
 Db 2119 CATAAATCAACAGCTCACAATTTGAAGGAATCGAAAGAAAGAAAGATTAGCAAAATCAA 2178
 QY 644 LysGlnLys-----GluGluAlaGlnArgAla---GlnGluArgAspLysGlnTrp 660
 Db 2179 AAAAGAAACTAGAGATGAGCTGCAGGAAAGCAAGCAAGCAAGAAAGAAAGAAACTTGTGG 2238

QY 661 LeuGluHisValGlnGlnGluAspGluHisGlnArgProArgLysLeuHisGluGluGlu 680
 Db 2239 AGAGAAAGTATTAGAAAGGAAGAGAG-----GAAAGCAA 2274
 QY 681 LysLeuLysArgGluGluSerValLysLysLysAspGlyGluGluLysGlyLysGlnGlu 700
 Db 2275 AAACGACTCCAGCAAGAAAGAGTCAACAGCAAAACTCAAGAGAGAGCAAGAAAGCTGAG 2334
 QY 701 AlaGlnAspLysLeuGlyArgLeuPheHisGlnHisGlnGluProAlaLysProAlaVal 720
 Db 2335 GCA----- 2337
 QY 721 GlnAlaProTrpThrAlaGluLysGlyProLeuThrIleSerAlaGlnGluAsnVal 740
 Db 2338 -----AAACAAAGTGAGACAGCCAGTCT----- 2361
 QY 741 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 760
 Db 2362 ---TTGCTGAATTACAGAGCACTGTACCTTTTGAAGCAAGAAACCATGATGATGATGAT 2418
 QY 761 IleGlnProGlyAspIleVal 767
 Db 2419 TTTAGTTCTCGGATATATTT 2439
 RESULT 5
 AK034367 2352 bp mRNA linear HTC 18-SEP-2003
 LOCUS Mus musculus adult male dienecephalon cDNA, RIKEN full-length
 DEFINITION enriched library, clone:9330184B19 product:SH3 domain protein 1B,
 full insert sequence.
 ACCESSION AK034367
 VERSION AK034367.1 GI:26083928
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Mech. Enzymol. 303, 19-44 (1999)
 99279253
 PUBMED 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 PUBMED 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 PUBMED 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.

TITLE
JOURNAL
REFERENCE
AUTHORS

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2352)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 URL: http://location.qualifiers

FEATURES
 source

1..2352
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:9330184B19"
 /db_xref="MGI:2398567"
 /db_xref="taxon:10090"
 /clone="9330184B19"
 /sex="male"
 /tissue_type="diencephalon"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 misc_feature
 1..2352
 /note="SH3 domain protein 1B (MGD|MGI:1338049)"

ORIGIN

Alignment Scores:
 Pred. No.: 6,97e-96 Length: 2352
 Score: 1645.50 Matches: 364
 Percent Similarity: 70.85% Conservative: 93
 Best Local Similarity: 56.43% Mismatches: 128
 Query Match: 27.91% Indels: 60
 DB: 11 Gaps: 18

US-09-720-934-2 (1-1143) x AK034367 (1-2352)

QY 1 MetAlaGlnPheProThrProPheGlySerLeuAspIleThrAlaIleThrValGlu 20
 Db 262 ATGGCTCAGTTTCCACAGCGATGAATGGAGGGCCAAATATGTGGGCTATTACATCTGAA 321
 QY 21 GluArgAlaLysHisAspGlnGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
 Db 322 GAACGTACTAAGCATGATAACAGTTTGATGAACCTCAACCTTCAGAGGTTACATAACA 381
 QY 41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
 Db 382 GGTGATCAAGCCCGTACTTTTCTACAGTCAGTCTGCGCGCCCGGTTTACGTGAA 441
 QY 61 IleThrAlaLeuAlaAspMetAsnAspGlyArgMetAspGlnValGluPheSerIle 80

442 ATATGGGCGCTTATCAGATCTGAACAGGATGGGAAGATGACAGCAGCAAGTTCTCTATA 501
 QY 81 AlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
 Db 502 GCTATGAACATCATCAAGTTAAAGTTGAGGGCCCAACAGCTGCTGTAGTCTCTCCCTCT 561
 QY 101 ValMetLysGlnGlnProVal-----AlaIleSerSerAlaProAlaPheGlyMetGly 118
 Db 562 ATCATGAAACACACCCCTATGTTCTCTCCACTAATCTGTCT---CGTTTGGGATGGGA 618
 QY 119 GlyIleAlaSerMetPro-----ProLeuThrAlaValAlaProVal 132
 Db 619 -----AGATGCCCAATCTGTCTCATTCATCAGCCATGTGCTCCAGTTGCACCTATA 669
 QY 133 -----ProMetGly-----SerIleProValValGlyMetSerPro 144
 Db 670 GCAACACCCCTGTCTCTGCTACTTCCAGGGACCCAGTATTCTCCCTTAATGATGCTGCT 729
 QY 145 ThrLeuValSerSerValProThrAlaAlaValProLeuAlaAlaGlnGlyAlaProPro 164
 Db 730 CCCTAGTGCCTTCTGTAGTACATCTCATTTACCA-----AATGGAATGCCAGT 780
 QY 165 ValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuProLysSerSerSer 184
 Db 781 CTCATTGAGCCCTTA---TCCATTCCTTATCTCTTCACATGCTCTCATGATCATCT 837
 QY 185 PheSer-----ArgSerGlyProGlySerGlnLeuAsnThrLysLeuLysAlaGln 202
 Db 838 TACAGCCTGATGATGGAGGATTGCT-----GGTGTAGTATCCAGAAAGGCCAG 888
 QY 203 Ser---PheAspValAlaSerValProPro----- 211
 Db 889 TCTCTGATTGATTGAGGATCTAGTACTCAACTTCCCTCACTTCCCTCTCAGGGAAC 948
 QY 212 -----ValAlaGluTyrAlaValProGlnSerSerArgLeuLysTyrArg 226
 Db 949 TCACCTAAGCAGGGACCTCAGAGTGGCGAGTTCTCCTCAGCTTCAAGATTAAAGTATCGG 1008
 QY 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
 Db 1009 CAAAAATTAATAGTTAGTACAAAGCATGAGCGGATACCTCTCAGGTTTTCAGGTAGA 1068
 QY 247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleThrAsnLeuSer 266
 Db 1069 AATGCCCTTCTTCTCAGTCAAAATCTCTCAAACTCAGTACTACTATTTCGACTCTGGCT 1128
 QY 267 AspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeuAlaMetHisLeuIle 286
 Db 1129 GACATCGATGCTACCGACAGTTGAAGCTGAAGATTATTCTGGCGATGCACTCACT 1188
 QY 287 AspValAlaMetSerGlyGlnProLeuProProValLeuProProGluTyrIleProPro 306
 Db 1189 GACATGGCCAAAGCTGGACAGCCACTACCACTCAGCTTGCTCCCGAGCTGTCCCTCCA 1248
 QY 307 SerPheArgArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326
 Db 1249 TCTTTC-----AGAGGGGAAAGCAAGTTGATTCTGTATTAATGGAAT----- 1290
 QY 327 GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnGlnLeuGluLysLysLeu 346
 Db 1291 -----CTGCCTCATATCAGAAAAACAGAGAGAGAGAGCTCAG-----AGAAACTG 1338
 QY 347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGlyAsnLeuGluLeuGlu 366
 Db 1339 CCAGTTACTTTTTCAGGACAAAACGAAAGCCCACTATGAACGAGGAAACATGAGCTGGAG 1398
 QY 367 LysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGlnGluArgLeuAlaGlnLeu 386
 Db 1399 AAGCGAGCCCAAGTGTGTGAGCAGCAGCAGCAGGAGGAGCTGAACGAAAGCCAGAAA 1458
 QY 387 GluArgAlaGluGlnGluArgLysGluArgGluArgGlnGlnGluArgLysArgGln 406
 Db 1459 GAGAGAGAGAGTGGGAGCGGAAACAGAGAACTCAAGAGCAAGATCGAAGAGCAG 1518

```

QY 407 LeuGluLeuGluLysGlnLeuGluLysGlnArgGlnLeuGluArgGlnGluGlu 426
Db 1519 CTGAGTTGGAGAAACCTTTGGAGAAACAGAGAGAGCTGGAGACAGCGGGAGAAAG 1578
QY 427 ArgArgLysGluLeuGluArgGluGluAlaLysArgGluLeuGluArgGlnArgGln 446
Db 1579 AGGAAAGAGAGATAGAAAGACGAGAGCGAGCAAAACAGGAGCTTGGAGACACGCGT 1638
QY 447 LeuGluTrpGluArgAsnArgGlnGlnLeuLeuLeuLeuLeuLeuLeuLeuLeu 466
Db 1639 TTAGATGGGAAAGACTCCGTGGCAGAGCTGCTCAGTCAGACAGACACGAGGAAACAGAA 1698
QY 467 AspIleValValLeuLysAlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsn 486
Db 1699 GACATTGTCAGGCTGAGCTCCAGAAAGAAAGTCTCCACCTGGAACTGGAGCAGTGAAT 1758
QY 487 AspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArgCysArgLeuThrThrGln 506
Db 1759 GGAAACATCAGCAGATCTCAGGAGACTACAGATGTCCAAATCAGAAAGCAACACAA 1818
QY 507 ArgGlnGlnLeuSerThrAsnLysSerArgGluLeuArgIleAlaGluLeuThrHi 526
Db 1819 AAGAGCTGAGCTAGAGTTTGGATAAAACAGTGTGACCTGGAAATTTAGAAATCAAAACA 1878
QY 526 sLeuGlnGlnLeuGlnSerGlnGlnMetLeuGlyArgLeuLeuProGluLysG1 546
Db 1879 ACTTCAACAGAGCTTAAAGGAATATCAAAATAAAGCTTATCTGTGCTCCAGAGACA 1938
QY 546 nIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsnSerLeuHisArgSerLeuVa 566
Db 1939 GCTATTAAACGAAGATTAATAACATGCAGCTCAGTAACACA---CCTGATTCAGGAT 1995
QY 566 lThrLeu-----LysArgAlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAs 584
Db 1996 CAGTTTACTTCATAAAAAAGTCATCAGAAAGAGAAATTAATGCCAAAGACAT---AAAGA 2052
QY 584 pGlnLeuAspGluValGluLysGluThrArgSerLysLeuGlnGluLeuAspIlePheAs 604
Db 2053 ACAATTAGATGCTCTGAAAAGAAACTGCATTAAGCTCTCAGAAATGGATTCATTAA 2112
QY 604 nAsnGlnLeuLys 608
Db 2113 CAATCAGCTGAAG 2125

RESULT 6
LOCUS BQ719508
DEFINITION AGENCOURT 8219793 Lupski_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6188111 5', mRNA sequence.
ACCESSION BQ719508
VERSION BQ719508.1 GI:21858405
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 979)
NH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAW13583 row: c column: 24
High quality sequence stop: 643.
Location/Qualifiers

ORIGIN
Alignment Scores:
Pred. No.: 2,39e-82 Length: 979
Score: 1428.50 Matches: 296
Percent Similarity: 95.25% Conservative: 5
Best Local Similarity: 93.67% Mismatches: 12
Query Match: 24.23% Indels: 4
DB: 13 Gaps: 1
US-09-720-934-2 (1-1143) x BQ719508 (1-979)
QY 98 LeuProValMetLysGlnProValAlaIleSerSerAlaProAlaPheGlyMet 117
Db 2 CTCCCCCTGTGATGAACACCAACAGTGTCTATTTCTAGCCACAGCAATTTGGTATG 61
QY 118 GlyGlyIleAlaSerMetProLeuThrAlaValAlaProValProMetGlySerIle 137
Db 62 GGAGTATCGCCAGCATGCCACCGCTTACAGCTGTGCTCCAGTGCCAATGGATTCATT 121
QY 138 ProValValGlyMetSerProThrLeuValSerSerValProThrAlaAlaValProPro 157
Db 122 CCAGTTGTGGAAATGTCTCCAACCTAGTATCTTGTCTCCACAGCAGCTGTGCCCCC 181
QY 158 LeuAlaAsnGlyAlaProValIleGlnProLeuProAlaPheAlaHisProAlaAla 177
Db 182 CTGCTTAACGGGGCTCCCTCTGTATCAACCTCTGCTGCTGATTTGCTATCTCCGACCC 241
QY 178 ThrLeuProLysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrIys 197
Db 242 ACATTGCCAAGAGTCTCTCTCTTTAGTAGATCTGTCAGGGTCAACACTAAACACTAAA 301
QY 198 LeuGlnLysAlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaVal 217
Db 302 TTCAAAAGGCACAGTCAATTGTGTGGCCAGTGTCCACAGTGGCAGAGTGGGCTGTT 361
QY 218 ProGlnSerSerArgLeuLysTyArgGlnLeuPheAsnSerHisAspLysThrMetSer 237
Db 362 CCTCAGTCATCAAGACTGAATACAGCAATATTCAATAGTATGATGACAAAATATGAT 421
QY 238 GlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAla 257
Db 422 GGACACTTAACAGGTCCTCCCAAGCAAGAACTATTCTTATGCTAGTCAAGTTTACCAAGCT 481
QY 258 GlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGlu 277
Db 482 CAGCTGGCTTCAATGATGGAAATCTTCTGACATGATCAAGTAAGAAATCTTACACAGAG 541
QY 278 GluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProPro 297
Db 542 GAATTTATCTTGGCAATGCACTCATTTAGTATGATGTCTGGCAACCACTGCCACCT 601
QY 298 ValLeuProGluTyIleProSerPheArgValArgSerGlySerGlyIle 317

```

```

Db      602 GTCCTGCCCTCCAGAAACATCCATCTCTTTAGAGAGATTCGATCTGCAGTGTATA 661
QY      318 SerValIleSerThrSerValAspGlnArgLeuProGluGluProValLeuGluAsp 337
Db      662 TCTGTCTAAGCTCAACATCTGTAGATCAGAGCTACAGAGAAACAGTTTAGAAGAT 721
QY      338 GluGlnGlnGlnLeuGluLysLysLeuProValThrPheGluAsp-LysLysArgGluAs 357
Db      722 GAACAACAACAATTAGAAAAGAAATTTACCTGTACGCTTTGAAGATAAAGAGCGGAGAA 781
QY      357 nPheGluArgGlyAsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnArg 377
Db      782 CTTTGAACGTGGCAACTGGAACTGGAGAAACGAAG-CAAGCTCTCTCTGNAACGCGAGCG 840
QY      377 glysGluGlnGluArgLeuAlaGlnLeuGluArgAlaGluGlnGlu-ArgLysGluArgG 397
Db      841 CAAAGGAAGCAGAGCGCTGGCCCANCTGAACGGCGCCAGCAGAAAGAAAGAAAGACGTG 900
QY      397 luArgGlnGluGlnGluArgLysArgGlnLeuGluLeuGluLys 411
Db      901 ACCCCAGACAGACAGCGCAAAAG---ACACTGGGACTGGAAAAA 941

RESULT 7
CF729634
LOCUS
DEFINITION
  UI-M-HD0-cld-n-08-0-UI.r1 NIH BMAP_HD0 Mus musculus cDNA clone
  IMAGE:30615007 5', mRNA sequence.
ACCESSION
  CF729634
VERSION
  CF729634.1 GI:37603802
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 873)
REFERENCE
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaabs-c@mail.nih.gov
  Tissue Procurement: Dr. James Lin University of Iowa
  CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
  CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Cloning Distribution: Dr. M. Bento Soares, University of Iowa
  Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/mousefl.html
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
Seq primer: pyX-5.
  Location/Qualifiers
    1..873
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6"
      /db_xref="taxon:10090"
      /clone="IMAGE:30615007"
      /tissue_type="whole eye"
      /dev_stage="embryo 12.5,13.5,14.5 dpc"
      /lab_host="DH10B (T1 phage resistant)"
      /clone_lib="NIH_BMAP_HD0"
      /notes="Organ: Eye; Vector: pyX-Asc; Site:1: EcoR I;
      Site:2: Not I; The library was constructed according
      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
      1996. Denatured RNA was size fractionated on a 1% agarose
      gel. First strand cDNA synthesis was primed with oligo-dT
      primer containing a Not I site. Double strand cDNA was
      size selected according to mRNA size fraction, ligated
      with EcoR I adaptor, digested with NotI and then cloned
      directionally into pyX-Asc vector. The library tag
      sequence located between the Not I site and the polyA tail
      is TTATTGAGT. This library was created for the University
      of Iowa Brain Anatomy Project (BMAP): Gene Discovery in the

```

```

ORIGIN
Alignment Scores:
  Pred. No.:      4,87e-73      Length:      873
  Score:          1283.50      Matches:    263
  Percent Similarity: 94.06%      Conservative: 6
  Best Local Similarity: 91.96%      Mismatches: 17
  Query Match:    21.77%      Indels:     3
  DB:             14          Gaps:        0

US-09-720-934-2 (1-1143) x CF729634 (1-873)
QY      1 MetAlaGlnPheProThrProPheGlyGlySerLeuAspIleTrpAlaIleThrValGlu 20
Db      6 ATGGCTCAGTTTCCCAACACTTTCGGTGTAGCTGGATGCTGGGCCATAACTGTGGAG 65
QY      21 GluArgAlaLysHisAspGlnPheHisSerLeuLysProIleSerGlyPheIleThr 40
Db      66 GAAAGGGCCCAAGCATACCAGCAGATTCTTAGCTGAAGCCGATAGCGGGATTTATTACT 125
QY      41 GlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuProGlnProValLeuAlaGln 60
Db      126 GGTGATCAAGCGAGGAACTTTTTCCTGCTGGGTACTCAGCTGCTTTAGCACAA 185
QY      61 IleTrpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAspGlnValGluPheSerIle 80
Db      186 ATATGGGCGCTAGCGGACATGAATAACGATGAAGGATGATCAAGTGAATTTTCATA 245
QY      81 AlaMetLysLeuLysLysLeuLysLeuGlnGlyTyrGlnLeuProSerAlaLeuProPro 100
Db      246 GCCATGAAGCTTATCAAACTGAAGCTACAGGATATCAGCTCCCTCCACACTTCCOCT 305
QY      101 ValMetLysGlnGlnProValAlaIleSerSerAlaProAlaPheGlyMetGlyGlyIle 120
Db      306 GTCATGAACAGCAACAGTGGCTATTTCAGTGCACCACTTTGGTATAGAGGGGATT 365
QY      121 AlaSerMetProProLeuThrAlaValAlaProValProMetGlySerIleProValVal 140
Db      366 GCTAGCATGCCACCACTCAAGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
QY      141 GlyMetSerProThrLeuValSerSerValProThrAlaAlaValProProLeuAlaAsn 160
Db      426 GGAATGTCTCCACCTCTAGTATCTTCTGCTCCCTCCAGCAGCAGTGCCTCCCTGGCTAAC 485
QY      161 GlyAlaProProValIleGlnProLeuProAlaPheAlaHisProAlaAlaThrLeuPro 180
Db      486 GGGGCTCTCTCCCTCATACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
QY      181 LysSerSerSerPheSerArgSerGlyProGlySerGlnLeuAsnThrLysLeuGlnLys 200
Db      546 AAGAGTTCTTCCCTTCAGCAGATCTGTCAGGGTCAAAATTAACACTAAGATTACAGAAG 605
QY      201 AlaGlnSerPheAspValAlaSerValProProValAlaGluTrpAlaValProGlnSer 220
Db      606 GCACAATCATTCGATGTCGCCGCGCCCTCCAGCAGCAGAGATGGGCTGTGCTCAGTCA 665
QY      221 SerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeu 240
Db      666 TCANGGCTGAATACAGGCAGTTATTTCACAGCCAGCAGACANAACTATGAGTGCACACTTA 725
QY      241 ThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAla 260
Db      726 ACAGGTCNCCAGGCAAG-ACTATTCTCATGCANTCANGTTTACCCCGAGGCTCAGCTGGCT 784
QY      261 SerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIle 280
Db      785 TCANNTATGGATCTTTCTGACATTTGATCAAGATGGAAA-CTCACTGCAGGAGAATT-ATC 842
QY      281 LeuAlaMetHisLeuIle 286
Db      843 CTAGCTATGAACCTATTG 860

```

Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

```

RESULT 8
BQ942708
LOCUS
DEFINITION BQ942708 911 bp mRNA linear EST 21-AUG-2002
5', mRNA sequence.
ACCESSION BQ942708
VERSION BQ942708.1 GI:22358186
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 911)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD
CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM3743 row: n column: 03
High quality sequence start: 10
High quality sequence stop: 564.
Location/Qualifiers
1. 911
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6314690"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_129"
/notes="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
Site1: EcoRV, Site 2: NotI; cloned unidirectionally.
Primer: Oligo dT. Average insert size 2.2 kb. Constructed
by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
Library."
ORIGIN
Alignment Scores:
Pred. No.: 4,14e-72 Length: 911
Score: 1269.50 Matches: 280
Percent Similarity: 93.83% Conservative: 9
Best Local Similarity: 90.91% Mismatches: 10
Query Match: 21.54% Indels: 10
DB: 13 Gaps: 1
US-09-720-934-2 (1-1143) x BQ942708 (1-911)
QY 227 GlnLeuPheAsnSerHisAspLysThrMetSerGlyHisLeuThrGlyProGlnAlaArg 246
Db 2 CAGTTATTCACAGCCAGCAGCAAACTAGTGGACACTTACAGGTCCTCCAGGCAAGA 61
QY 247 ThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSer 266
Db 62 ACTATTCTCATGCAATCAAGTTTACCCAGGCTCAGCTGGCTTCAATATGGAATCTTCT 121
QY 267 AspIleAspGlnAspGlyLysLeuThrAlaGluPheIleLeuAlaMetHisLeuIle 286
Db 122 GACATTGATCAAGATGGAAACTACTGCAAGAAATT-ATCCTAGCTATGCACCTAATT 180
QY 287 AspValAlaMetSerGlyGlnProLeuProValLeuProGluTyrIleProPro 306
Db 181 GATGTTGCATGTCTGGTCAGCCACTGCGCGCGCTGCTCCCTCCAGAAATACATCCCTCT 240
QY 307 SerPheArgValArgSerGlySerGlyIleSerValIleSerSerThrSerValAsp 326
Db 241 TCCTTCAGAGAGTTCGTCGGCAGTGGGATCGCGTCAATAGCTCTTCTTGTGGAT 300

```

```

QY 327 GlnArgLeuProGluGluProValLeuGluAspGluGlnGlnLeuGluLysLysLeu 346
Db 301 CAGAGGCTGCTCAGAGAGCCGCTGCTCAGAGATGAGCAGCAGCCA---GAGAGAAACTG 357
QY 347 ProValThrPheGluAspLysLysArgGluAsnPheGluArgGlyAsnLeuGluLeuGlu 366
Db 358 CCTGTGACATTTGAAGATAAGAAAGCGGAGAACTTCGAGCGAGGACAGTGTGGAGCTGG 417
QY 367 LysArgArgGlnAlaLeuLeuGlnGlnArgLysGlnGlnGlnGlnGlnGlnGlnLeu 386
Db 418 AAGCGCGCCCAAGCGCTCTTGGAGCAGCAGCGCAAGAGCAGGAGCGGTGGCTCAGCTG 477
QY 387 GluArgAlaGlnGlnGluArgLysGluArgGluArgGlnGlnGlnGluArgLysArgGln 406
Db 478 GAGCGCCCGCAGCAGGAGAGAGCGGAGCGCGCAGCAGCAGGAGCGCAGCGGAG 537
QY 407 LeuGluLeuGluLysGlnLeuGluLysGlnArgGluLeuGluArgGlnArgGluGlu 426
Db 538 CTGGAGCTGGAGAAAGCAGCTGGAGAACGAGCGGAGCTGGAGCGGCGCAGCAGGAGAG 597
QY 427 ArgArgLysGlnIleGluArgArgGluAlaAlaLysArgGluLeuGluArgGlnArgGln 446
Db 598 AGGAGGAGAGAGATCGAGAGCGCGCGCGCAAAACGGGAACCTGGAAAGCGCAGCA 657
QY 447 LeuGluTrpGluArg-AsnArgArgGlnGlnLeuLeuAsnGlnArgAsnLysGluGln 466
Db 658 CTTGATGGGAACCGGNAACCGGACAGCAGGAACCTCTGAATCAGAGGAACAGGAGCAG 717
QY 466 u-AspIleValValLeuLysAlaLysLysThrLeuGluPheGluLeuGluAlaLeu 485
Db 718 GGGCCACCCCTGGTCTCTGAAAGCAAGGAGGAAAGACTCTGGAGTTTGAAGTACAACT 777
QY 486 AsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAspIleArgCysArg-LeuThrTh 505
Db 778 AATGACAAAAGCATCAGTAGAAGGAAAACCTCCGGATATCATCGTGGTCCACTGGCAAC 837
QY 505 rGlnArg-GlnGluIleGluSerThr-AsnLysSer-ArgGluLeuArgIleAla-GluI 524
Db 838 CCAGAGGCGCAAGAAATTGAGGACGACCAACAGTCTAAGAGGAGCTAAATTTGTTGAA 897
QY 524 leThrHis 526
Db 898 TCACCCAC 905
CA324450 818 bp mRNA linear EST 09-JUL-2003
UI-M-FY0-ccl-a-05-0-UI.r1 NIH BMAP_FY0 Mus musculus cDNA clone
IMAGE:6820974 5', mRNA sequence.
ACCESSION CA324450
VERSION CA324450.1 GI:24542548
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 818)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA

```

sequence: 338-455, >(GGA)n#Simple_repeat
Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers

1. .818
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6820974"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGACACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1.02e-71 Length: 818
Score: 1262.50 Matches: 256
Percent Similarity: 96.34% Conservative: 7
Best Local Similarity: 93.77% Mismatches: 9
Query Match: 21.42% Indels: 1
DB: 14 Gaps: 1

US-09-720-934-2 (1-1143) x CA324450 (1-818)

QY 277 GUGLUPHETLEULALAWETHLSLEULASPVALALAMETSERGLYGLNPROLEUPRO 296
DB 2 GAAGAATTATCTAGTATGACCACTAATGATGTGCCAIGTCTGTCCAGCCAGTCCG 61
QY 297 PROVALLEUPROGLUTYRILLEPROPROSERPHEARGARGVALARGSERGLYSERGLY 316
DB 62 CCGTCTCGCTCCAGAAATACATCCCTCTCTTCAGAAAGATTCGCTCCGCGAGTGG 121
QY 317 ILESERVALLLESERSETHRSERVALASPLNARGLEUPROGLUINPROVALLEUGLU 336
DB 122 ATGTCCGTCATAGCTCTCTCTGTGTGATCAGAGGCTGCTCAGGAGCGCTCGTCAGAG 181
QY 337 ASPGLUGLNGLNLNLEULYLSLYSLEUPROVALTHRPHEGLUASPLYSLYSARGGLU 356
DB 182 GATGACGACGAGCCCA---GAGAGAACAACCTGCTGTGACATTTGAGATAGAGAGCGGAG 238
QY 357 ASPNHEGLUARGGLYASNLEUGLULEULYLSARGARGGLNLALEULEUGLNGLNLN 376
DB 239 AACTTCGAGCGAGCGAGTGTGGAGCTGAGAACGCCCGCCCAAGCGCTCTTGGAGCAG 298
QY 377 ARGLYSGLUGLNLGUAARGLEUALAGLNLNLEULUARGNLAGLUGLNLGUAARGLYSGLUARG 396
DB 299 CGCAAAAGACGAGCGCGTGTGGCTCAGCTGAGCGCGCCGAGCAGGAGAGAAAGACGG 358
QY 397 GLUARGGLNGLNLGUAARGLYSARGGLNLEULUARGGLYSGLNLNLEULYSGLN 416
DB 359 GAGCGCCAGAGGACGAGCCAGCGCACCTGGAGCTGGAGAGCAGCTGGAGAGCAG 418
QY 417 ARGGLULEUGLUARGGLNARGGLUGLUARGGLYSGLULEGLUARGGLUALA 436
DB 419 CGGAGCTGAGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGATCGAGAGCGCGAGGCC 478
QY 437 ALALYSARGGLULEUGLUARGGLNARGGLNLEUGLUARGGLUARGGLNLEUGLU 456

DB 479 GCAGAAACGGGAACTGGAAAGCGCAGCACTTGAATGGAAACGGACGAGACAGGAA 538
QY 457 LEULEUASNGLNARGASNLYSGLUGLNLGUAASPILEVALLEULYSLALALYSLYSLYS 476
DB 539 CTCTGTAATCAGAGGAACAAGGAGCAGGAGGCGCACCGTGTGCTGAAGCGAGGAGGAG 598
QY 477 THRLEUGLUPHEGLULEUGLUALAULEUASNASPLYSGLYSHISGLNLNLEUGLUGLY 496
DB 599 ACTCTGGAGTTTGAGTTAGAGCTCTGAATGACAAAAGCATCAGCTAGAGGAGAAACTT 658
QY 497 GINASPILLEARGCYSGARGLEUTHRGLNARGGLNLEUGLULEGLUSERTHRASNLYS 516
DB 659 CAGGATATCAGGTGTGCTGCTGCAACCCAGAGGCAAGAAATTCAGAGCAGCAAGATCT 718
QY 517 ARGGLULEUARGILEALAGLUILERTHRHISLEUGLNLGGLNLEUGLNUSERGGLN 536
DB 719 AGAGAGCTAGAAATTCCTGANATCACCACCTACAGCAGCAGTTGCAGGATTCAGCAN 778
QY 537 METLEUGLYARGLEULEULEPROGLULYSGLNILELEUASN 549
DB 779 ATGCTTGGNAGACTTATTCCAGAGANACAGATACTCAGT 817
RESULT 10
BU319490
LOCUS
DEFINITION
BU319490 886 bp mRNA linear EST 28-NOV-2002
603485809F1 CSRQCHN62 Gallus gallus cdna clone CHEST380a4 5', mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
FEATURES
source
1. .886
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST380a4"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN62"
/note="Organ: heads; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adaptors, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Donaldo et al., Genome Research 6

(1996): 791, except that a significantly longer reamealing hybridization was used."

ORIGIN		Alignment Scores:		Length:	
Pred. No.:	2,56e-71	1257.00	268	Matches:	886
Percent Similarity:	94.63%	Conservative:	14	Indels:	10
Best Local Similarity:	89.93%	Query Match:	6	Gaps:	2
DB:	13				
US-09-720-934-2 (1-1143) x BU319490 (1-886)					
Qy	374	GlulGlnArgLysGluGln-GluArgLeuAlaGlnLeuGluArgAlaGluGlnGluAr	393		
Db	2	GAACAGCAACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	61		
Qy	393	gLySgLuArGluArgGlnGlnGlnGluArgLysArgGlnLeuGluLysGlnLe	413		
Db	62	GAAGCAAGCTGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTT	121		
Qy	413	uGlLysGlnArgGluLeuGluArgGlnGluGlu-GluArgLysGlnLeuGluA	433		
Db	122	GGAGAAACAACGGCAATTGGAAACGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	181		
Qy	433	rgArgGluAlaAlaLysArgGluLeuGluArgGlnArgGlnGlnGluArgGlnGln	453		
Db	182	GAAGAGAGGCTGCAAAACGGCAAGCTTGGAGGCAAGCAAGCAAGCAAGCAAGCAAG	241		
Qy	453	rgArgGlnGluLeuLeuAsnGlnArgAsnLysGlnGlnGlnGlnGlnGlnGln	473		
Db	242	GTCGGCAGAACTCTTAATCAAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG	301		
Qy	473	laLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeu	493		
Db	302	CAAGAAGAGAGACTTGAATTTGAGCTGGAGGCTCTAAATGATAAATAAATAAATCAGT	361		
Qy	493	luGlyLysLeuGlnAspLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	513		
Db	362	AAGGAAAGCTTCAGGATATCAGATGTCGGCTGTCTACCAAGCAAGCAAGCAAGCAAG	421		
Qy	513	hrAsnLysSerArgGluLeuArgLysAlaGluLysLysHisGlnGlnGlnGlnGlnGln	533		
Db	422	CAATAAATCTAGAGAACTGAGATTTGAGAAATCACCAATTTGCAACAGCAGCTACAG	481		
Qy	533	luSerGlnGlnMetLeuGlyArgLysLysProGluLysGlnLysLeuAsnAspGlnLeuL	553		
Db	482	AGTCTCAGCAGATGCTTGGAAAGTTGATTCAGAAAGCAATTAATCAATGACCAACTAA	541		
Qy	553	ysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuG	573		
Db	542	AGCAGGTTCAACAGAACAGTTTGCATAGAGATTCCTTCTTACTATCAAAAGAGCCTTAG	601		
Qy	573	luAlaLysGluLeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluT	593		
Db	602	AAGCAAGAACTAGCCCGTCAACAGCTTCAGAGCCAGCTAGATGAGTGAAGCAAGAA	661		
Qy	593	hrArgSerLysLeuGlnGlnLysLysLysLysLysLysLysLysLysLysLysLysLys	613		
Db	662	CCAGATCTAAATCTCAGGAAATCGATATTTTCAATAATCAGCTGAAGAGCTGAGAGAA	721		
Qy	613	leHisAsnLysGlnGlnLeuGlnLysLysLysLysLysLysLysLysLysLysLysLys	632		
Db	722	TACATAACAGACAGCAGCTTCAGAAACAAAGAAACCTTGAAGCTGAGAGCTGAAACA	781		
Qy	632	nLysGluGlnGluArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys	652		
Db	782	GAAGAAGCAAGAAAGG---ACGACAGAACTGGAAAGCAAGAA---GAAGCTCAAGGCG	835		
Qy	652	gAlaGlnGluArgAspLysGlnTrpLeuGluHisValGlnGlnGlnGlu 667			
Db	836	AATCCAGGATCGGGATAAACAACCGTTGATCCGAGTGCACCAAGAG 881			

RESULT 11 LOCUS

BM681943 773 bp mRNA linear EST 27-FEB-2002
UI-E-E01-aiw-e-23-0-UI.s1 UI-E-E01 Homo sapiens cDNA clone
UI-E-E01-aiw-e-23-0-UI 3', mRNA sequence.

ACCESSION
BM681943

VERSION
BM681943.1 GI:18991839

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 773)

AUTHORS
Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE
Normalization and subtraction: two approaches to facilitate gene

JOURNAL
discovery

MEDLINE
97044477

PUBMED
8889548

COMMENT
Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA

sequence: 1-53, >POLY A#Simple_repeat (matched complement)

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1..773

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="UI-E-E01-aiw-e-23-0-UI"

/tissue_type="fetal eye"

/dev_stage="fetal"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-E01"

/note="Organ: eye; Vector: p773-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I;

UI-E-E01 is a normalized cDNA library containing the

following tissue(s): fetal eye. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into p773-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

CGCGTATACC. This library was created for the program, Gene

Discovery in the Visual System, supported by National Eye

Institute (NEI).

TAG_TISSUE=human fetal eye

TAG_LIB=UI-E-E01

TAG_SEQ=CGCGTATACC"

ORIGIN

Alignment Scores:

Pred. No.: 2e-70

Score: 1242.00

Percent Similarity: 99.22%

Best Local Similarity: 98.05%

Length: 773

Matches: 252

Conservative: 3

Mismatches: 2

```
Query Match: 21.07% Indels: 0
DB: 12 Gaps: 0
US-09-720-934-2 (1-1143) x BM681943 (1-773)

Qy 391 GlnGluArgLysGluArgGlnGluArgGlnGluArgLysArgGlnLeuGluGlu 410
|||
Db 773 CAGAGAGGAGGAGCGTGCAGCCCGAGGAGCAAGAGCGCAAAAGACAACCTGGAACTGGAG 714
|||
Qy 411 LysGlnLeuGluLysGlnArgGluLeuGluArgGlnArgGluGluArgArgLysGlu 430
|||
Db 713 AGCACTGGAAAGACAGCGGAGCTAGAACCGCAGAGAGATGAGAGGAGGAAAGAA 654
|||
Qy 431 IleGluArgArgGluAlaAlaLysArgGluLeuGluArgGlnArgGlnLeuGluTrpGlu 450
|||
Db 653 ATTGAGAGCGAGAGGCTGCAAAACGGAACTTTGAAAGGCAACGACAACTTGAGTGGAA 594
|||
Qy 451 ArgAsnArgGlnGlnLeuLeuLeuAsnGlnArgAsnLysGluGlnAspIleValVal 470
|||
Db 593 CGGAATCGAAGCGCAAGAACTTACTAAATCAAGAAACCAAGAAAGAGGACATAGTTGTA 534
|||
Qy 471 LeuLysAlaLysLysLysThrLeuGluPheGluLeuGluAlaLeuAsnAspLysLysHis 490
|||
Db 533 CTGAAGCAAAAGAAAGACTTTGGAAATTTGAATTAGAACTCTTAATGATAAAGCAT 474
|||
Qy 491 GlnLeuGluGlyLysLeuGlnAspIleArgCysArgLeuThrThrGlnArgGlnGluIle 510
|||
Db 473 CAACTAGAGGGAAACTTCAAGATATCAGATGTCGATTGACCAACCAAGGCAAGAAAT 414
|||
Qy 511 GluSerThrAsnLysSerArgGluLeuArgIleAlaGluIleThrHisLeuGlnGlnGln 530
|||
Db 413 GAGAGCACAAACAAATCTAGAGAGTTGAGAAATTCGCGAAATACCCATCTACAGCAACA 354
|||
Qy 531 LeuGlnGluSerGlnGlnMetLeuGlyArgLeuIleProGluLysGlnIleLeuAsnAsp 550
|||
Db 353 TTACAGGAATCTCAGCAAAATGCTTGGAGACTTATTCAGAAACACGATPACTCAATGAC 294
|||
Qy 551 GlnLeuLysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArg 570
|||
Db 293 CAATTAAACCAAGTTCAGCAGAACAGTTTCACAGAGATTCACTTGTGTACACTTAAAGA 234
|||
Qy 571 AlaLeuGluAlaLysGluLeuAlaArgGlnHisLeuArgAspGlnLeuAspGluValGlu 590
|||
Db 233 GCCTTAGAAGCAAAAGAACTAGCTCGCAGCACCTACGACCAACTGGATGAGTGGAG 174
|||
Qy 591 LysGluThrArgSerLysLeuGlnGluIleAspIlePheAsnAsnGlnLeuLysGluLeu 610
|||
Db 173 AAGAAACTAGATCAAAACTACAGGAGATTGATATTTCAATATCAGCTGAAGGAACTA 114
|||
Qy 611 ArgGluIleHisAsnLysGlnGlnLeuLysGlnLysSerMetGluAlaGluArgLeu 630
|||
Db 113 AGAGAAATACACAATAAGCAACAACTCCAGAAGCMAAAGTCCATGGAGGCTGAACGACTG 54
|||
Qy 631 LysGlnLysGluGlnGluArgLysIleIleGluLeuGlnLysGlnLysGlu 647
|||
Db 53 AAACAGAAAGCAAGAACGAAAGATCATAGATTAGAAAAAAGAAAAA 3

RESULT 12
BU317489
LOCUS BU317489 886 bp mRNA linear EST 28-NOV-2002
DEFINITION 603849519F1 CSEQCHN62 Gallus gallus cDNA clone CHEST843f5 5', mRNA
sequence.
ACCESSION BU317489
VERSION BU317489.1 GI:25825490
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 886)
AUTHORS Boardman,P.E., Sanz-Exquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Pong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
```

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source

1..886
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, HiseX"
/db_xref="taxon:9031"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN62"
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcORI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcORI, size-selected, and cloned into the NotI and EcORI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores: 4.67e-70 Length: 886
Pred. No.: 1237.50 Matches: 269
Score: 94.31% Conservative: 13
Percent Similarity: 89.97% Mismatches: 10
Best Local Similarity: 20.99% Indels: 8
Query Match: 13 Gaps: 1
DB: 1
US-09-720-934-2 (1-1143) x BU317489 (1-886)
Qy 242 GlyProGlnAlaArgThrIleLeuMetGlnSerSerLeuProGlnAlaGlnLeuAlaSer 261
|||
Db 3 GGTCCGCAAGCAAGAACTATTCTTATGCAGTCAAGTTTACCCAGGCTCAGCTGGCTACA 62
|||
Qy 262 IleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeuThrAlaGluGluPheIleLeu 281
|||
Db 63 ATATGGAATCTTTCAGATATTGATCAAGATGGAAGCTTACAGCTGAAGATTATTCTG 122
|||
Qy 282 AlaMetHisLeuIleAspValAlaMetSerGlyGlnProLeuProValLeuProPro 301
|||
Db 123 GCTATGCACTTAATTGATGTAGTATGTCTGTGTCAGCCACTGCCCTGTACTGCCTCCA 182
|||
Qy 302 GluTyrlleProProSerPheArgValArgSerGlySerGlyIleSerValIleSer 321
|||
Db 183 GAGTATATTCCACCTTCATTAGAGAGTACGCTCTGTGTAGTGTGTATCTGCTGAAGT 242
|||
Qy 322 SerThrSerValAspGlnArgLeuProGluGluProValLeuLeuAspGluGlnGln 341
|||
Db 243 TCAGTATCTGTAGCAACAAAGTTTACCAGAAAGAACCAACGATTAGAAAGAACAGACAA 302
|||
Qy 342 LeuGluLysLysLeuProValThrPheGluAspLysLysArgGluAsnPheGluArgGly 361
|||
Db 303 CTGGAAGAAAGAAATGCCAGTTACATTGAAGATAAAAGCTGAGAACTTTGAACGTGGC 362
|||
Qy 362 AsnLeuGluLeuGluLysArgArgGlnAlaLeuLeuGluGlnGlnArgLysGluGlnGlu 381
|||

Db	363	AACTAGAACTTGAARACGGAGCGAGCTCTCTCTGGAACAGCAACCAAGAACAGAG	422
Qy	382	ArgLeuAlaGlnLeuGluArgAlaGluGlnGluArgLysGluArgGlnGluGln	401
Db	423	CGTCTAGCAGCTGGAAACGGCAGAGCAAGAAAGGAAGAACTGTAACACAGAGCAA	482
Qy	402	GluArgLysArgGlnLeuGluGluGlnGluGlnGluGlnGluGlnGluGln	421
Db	483	GAAGAGAAAGACAACTGGAGCTAGAGAAACAGTTGGAGAAACACCGGAAATTTGAACGG	542
Qy	422	GlnArgGluGluGluArg---ArgLysGluLeuGlu-ArgArgGluAlaAlaLysArgGln	440
Db	543	CAGAGAAAGAGGACAGCGAGGACAGAAATAGAAACGAGAGAGCTGCAAAACGGGA	602
Qy	440	uLeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgGlnGlnLeuLeuAsnGln	460
Db	603	ACTTGAGAGGCAAAAGGCAACTTGAATGGGAACGTAATCGTCGGCAAGAACTGCTAAATCA	662
Qy	460	nArgAsnLysGluGlnGluAspLleValValLeuLys-AlaLysLysLysThrLeuGluP	480
Db	663	AAGAAACAGAGAAACAGAGGACATAGTTGTTCTGAGGGCAAGAGAGACTTTAGAAAT	722
Qy	480	heGluLeuGluAlaLeuAsnAsp-LysLysHisGlnLeuGluGlyLysLeuGlnAspLle	499
Db	723	TTGAGCTGGAGCTCTAGATGATCACACAAATCAGTTGGAGGAAGCTTCAGGATATC	782
Qy	500	ArgCysArgLeu-ThrThrGlnArgGlnGluGluLeuGluSerThrAsnLysSerArgGluLe	519
Db	783	AGATGTCGGCTGCTCTACCAAGACAGAAATTCAGAT-TCAAGTAAATCTAGAGAACT	841
Qy	519	uArg-LleAlaGlu-LleThrHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	533
Db	842	GAGAACTCGCAAGAAATACCCCTTTGGCAACAGCAGCTACAGGAC	886
RESULT 13			
BU441953		968 bp	mRNA
LOCUS	603208112F1	CSEQRBN11	Gallus gallus
DEFINITION			linear clone CHEST18393 5', mRNA
ACCESSION	BU441953		
VERSION	BU441953.1	GI:25931264	
KEYWORDS	EST		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
REFERENCE			
AUTHORS	Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E., Pong, W. T., Fickie, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.		
TITLE	A Comprehensive Collection of Chicken cDNAs		
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)		
MEDLINE	22335534		
PUBMED	12445392		
COMMENT	Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 0161208930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk.		
FEATURES			
source			
	Location/Qualifiers		
	1..968		
	/organism="Gallus gallus"		
	/mol_type="mRNA"		
	/strain="Layer and broiler"		
	/db_xref="taxon:9031"		
	/clone="CHEST18393"		
	/sex="Male and female"		
	/tissue_type="muscle"		
	/dev_stage="adult"		

/lab host="DH10B"
/clone lib="CSEQRBN11"
/note="Vector: pBluescript II KS(+); Site 1: EcoRI;
Site 2: NotI. This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN			
Alignment Scores:			
Pred. No.:	5.65e-70	Length:	968
Score:	1237.00	Matches:	271
Percent Similarity:	83.82%	Conservative:	14
Best Local Similarity:	79.71%	Mismatches:	19
Query Match:	20.98%	Indels:	36
DB:	13	Gaps:	3
US-09-720-934-2 (1-1143) x BU441953 (1-968)			
Qy	313	SerGlySerGlyLysLeuSerValLysSerThrSerValAspGlnArgLeuProGluGlu	332
Db	3	TCTGTTAGTGGTGTATCTGCTAAGTTCTAGTATCTGTAGACCAAGGTTTACCAAGAA	62
Qy	333	ProValLeuGluAspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	352
Db	63	CCAGCATTTAGAGAGAAAGAACAGCAGCACTGGAAAAGAAATTCACATTGAGAT	122
Qy	353	LysLysArgGluAsnPhenGlyAsnLeuGluGluGluGluGluGluGluGluGlu	372
Db	123	AAAAACGTTAGAACTTTGAACGTGGCAATCTAGAACTTGAAAAACGAGGAGCTCTC	182
Qy	373	LeuGluGlnGlnArgLysGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	392
Db	183	CTGGAACAGCAACGCAAGAGCAAGAGCGCTAGCACAGCTGGAAACGGGAGCAAGAA	242
Qy	393	ArgLysGluArgGluArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	412
Db	243	AGGAAGAACTGTAACACAGACAGCAAGAAAGAAAGAAAGACAGCTGGAGCTAGAGA	302
Qy	413	LeuGluLysGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGlu	432
Db	303	TTGAGAAACCAACGG-----	317
Qy	433	ArgArgGluAlaAlaLysArgGluGluGlnArgGlnLeuGluTrpGluArgAsn	452
Db	318	-----GAATGGAAAGCTAAT	332
Qy	453	ArgArgGlnGluLeuLeuAsnGlnArgAsnLysGluGlnGluAspLysValValLeu	472
Db	333	CGTGGCAAGAACTGCTAAATCAAGAAACAGACAGCAAGAGGACATATTGTTCTGA	392
Qy	473	AlaLysLysLysThrLeuGluPhenGluLeuGluAlaLeuAsnAspLysHisGlnLeu	492
Db	393	GCAAGAGAAAGAGACTTTAGAAATTTGAGCTGGAAGCTCTAAATGTATATAAAAAA	452
Qy	493	GluGlyLysLeuGlnAspLleArgCysArgLeuThrThrGlnArgGlnGluGluSer	512
Db	453	GAAGAAAGCTTCAGGATATCAGATCTCGCTGCTCTACCCAAAGACAGAAATTTGA	512
Qy	513	ThrAsnLysSerArgGluLeuArgLysAlaGluLysLeuGlnGlnGlnGlnGln	532
Db	513	ACAAATAAATCTAGAGAACTGAGAATTCAGAAATCACCCATTTCACACAGCAGCTAC	572
Qy	533	GluSerGlnGlnMetLeuGlyArgLeuLeuProGluLysGlnLleLeuAsnAspGlnLeu	552

```

Db      573  |||||GAGTCTCAGCAGATGCTTGGAAAGTTGATTCAGAAAGCAATTAATCAATGACCAACTA 632
QY      553  LysGlnValGlnGlnAsnSerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeu 572
Db      633  AAGCAGGTTCAACAGAACAGATTGTCATAGAGATCTCTTCTTACTATCAAAAGAGGCTTA 692
QY      573  GluAlaLysGluLeuAlaArgGlnHisLeuArgAspGlnLeuAspGlu-ValGluLysGI 592
Db      693  GAAGCCAAGGAACCTAGCCCGCTCAACAGCTTCGAGACGACGCTAGATGAAGTAGAAAAA 752
QY      592  uThrArgSer-LysLeuGlnGlnLeuAspLeuPheAsnAsn-GlnLeuLysGluLeu--- 610
Db      753  AACCATGCTTAAACCTTCAGGAATTTGATATTTTCAATTAATCAGCTGAAGGAGCTGGAG 812
QY      611  ArgGluIleHisAsnLysGlnGlnLeuGlnLys-GlnLysSerMetGluAlaGluArgLe 630
Db      813  AGAAATTACATTACAGACACGACGCTTCAGAAACCAAGACTTGGAAAGCTGGAAGGT 872
QY      630  uLysGlnLysGluGlnGluArgLysIleLeuGluLeuLysGlnLysGlu 647
Db      873  GAAACCGAAA---GGAACCAAGACCGCGGCGCAACTGGAAGGCAAGAA 921

```

RESULT 14

```

CF741757
LOCUS      819 bp      mRNA      linear      EST 10-OCT-2003
DEFINITION  UI-M-HB0-clj-f-03-0-UI.r1 NIH BMAP_HB0 Mus musculus cDNA clone
IMAGE:30617114 5', mRNA sequence.
ACCESSION  CF741757
VERSION    CF741757.1  GI:37638096
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE  1 (bases 1 to 819)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: Dr. James Lin University of Iowa
           cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
           DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
           Clone Distribution: Distribution information can be found at
           http://genome.uiowa.edu/distribution/mousefl.html
           This clone was contributed by the Brain Molecular Anatomy Project
           (BMAP).

```

FEATURES

```

source
Seq primer: pYX-5.
Location/Qualifiers
1..819
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30617114"
/tissue_type="whole eye"
/dev_stage="embryo 12.5, 13.5, 14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/note="Organ: Eye; Vector: pYX-Asc; Site:1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TTATTGAGT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

```

```

ORIGIN
Alignment Scores:
Pred. No.:      4, 92e-69      Length:      819
Score:          1221.00      Matches:    256
Percent Similarity: 94.53%      Conservative: 3
Best Local Similarity: 93.43%      Mismatches: 14
Query Match:    20.71%      Indels:     4
DB:             14           Gaps:       1

US-09-720-934-2 (1-1143) x CF741757 (1-819)
QY      75  GlnValGluPheSerIleAlaMetLysLeuLysLeuLysGlnGlyTyrGlnLeu 94
Db      3  CRAAGTGAATTTTCCATAGCATGAAGCTTATCAAACTGAAGCTACAAGGATATCAGCTC 62
QY      95  ProSerAlaLeuProProValMetLysGlnGlnProValAlaIleSerSerAlaProAla 114
Db      63  CCTCTCACACTTCCCTCTGTCATGAACACAGCAACCACTAGTGGCTATTTCCAGTGCACGCA 122
QY      115  PheGlyMetGlyGlyIleAlaSerMetProProLeuThrAlaValAlaProValProMet 134
Db      123  TTTGGTATAGGAGGAGATTGCTAGCATGCCACCACTCACAGCTGTTGCTCTGTGCCAATG 182
QY      135  GlySerIleProValValGlyMetSerProThrLeuValSerSerValProThrAlaAla 154
Db      183  GGCTCCATCCAGTGTGTGAATGTTCTCCACCTTAGTATCTTCTGCTCCAGCAGCA 242
QY      155  ValProProLeuAlaGlnGlyAlaProProValIleGlnProLeuProAlaPheAlaHis 174
Db      243  GTGCCTCCCTGGCTAACGGGGCTCTCCGTATACAGCTCTGCTGCTGCTGGTGGCAT 302
QY      175  ProAlaAlaThrLeuProLysSerSerSerPheSerArgSerGlyProGlySerGlnLeu 194
Db      303  CTTGCAGCCCAATGTCGAAGAGTTCTTCTTCCAGCAGATCTGCTCCAGGCTCACAAATTA 362
QY      195  AsnThrLysLeuGlnLysAlaGlnSerPheAspValAlaSerValProValAlaGlu 214
Db      363  AACACTAAGTTACAGAAAGGCACAACTCATTCGATGTCCAGCGCCCTCCAGCAGCAGAA 422
QY      215  TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 234
Db      423  TGGGCTGTGCTCAGTCATCAGAGCTGAAATACAGGCAGTATTTCACAGCCACGACAAA 482
QY      235  ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 254
Db      483  ACTATGAGTGGACCTTAACAGGTCCCAGGTCAGGCAAGAACTATTCTCATGCAATCAAGTTTA 542
QY      255  ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeu 274
Db      543  CCCAGGCTCAGCTGGCTTCAATATGAATCTTTTTCGACATTGATCAAGATGGAAGAACTC 602
QY      275  ThrAlaGluGluPheIleLeuAlaMetHisLeuLeuAspValAlaMetSerGlyGlnPro 294
Db      603  ACTGCAGAGAAGATTATTCCTAGCTATGACCTAAATGATGTGCCATGTCTGTGTCAGCCA 662
QY      295  LeuProProValLeuProProGluTyrIleProProSerPheArgArgValArgSerGly 314
Db      663  CTGCGCGCGCTCTGCTGCTCCAGAAATA-ATCCCTTCTTCTTCTTCCAGAGAGTTCGCTCGGC 721
QY      315  SerGlyIleSerValIleSerSerThrSerValAspGlnArgLeuProGluGluProVal 334
Db      722  AGTGGGATGTCCTGCANTAG-TCTTCTTCTGTCGATCAGAGGCTGCTGTAGGGAGCGCTCG 780
QY      335  LeuGluAspGluGlnGlnGlnLeuGluLysLysLeuProVal 348
Db      781  TCAGAGGATGAGCAGCAGCCA---GAGAGAA-CTGCTGTG 818

RESULT 15
CA448081/c
LOCUS      CA448081      751 bp      mRNA      linear      EST 08-NOV-2002

```

Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

DEFINITION UI-H-ED1-ayj-i-11-0-UI.s1 NCI CGAP ED1 Homo sapiens cDNA clone
 UI-H-ED1-ayj-i-11-0-UI 3', mRNA sequence.
 CA448081
 VERSION CA448081.1 GI:24812501
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 751)
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA sequence: 1-53, >POLY A#Simple_repeat (matched complement)
 Seq primer: M13 FORWARD
 POLYA-Yes

FEATURES

source Location/Qualifiers
 1..751
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-ED1-ayj-i-11-0-UI"
 /tissue_type="Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP ED1"
 /note="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line C8S. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
 TAG_TISSUE=chondrosarcoma
 TAG_LIB=UI-H-ED1
 TAG_SEQ=CGTCAAGGCT"

ORIGIN

Alignment Scores:
 Pred. No.: 7,97e-69 Length: 751
 Score: 1217.00 Matches: 247
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 99.20% Mismatches: 0
 Query Match: 20.64% Indels: 0
 DB: 14 Gaps: 0
 US-09-720-934-2 (1-1143) x CA448081 (1-751)
 Qy 399 GlnGluGlnArgLysArgGlnLeuGluLysGlnLeuGluLysGlnArgGlu 418
 Db 749 CAGGAGCAAGCGCAAAAGACAACTGCAACTGGAGAGCAACTGCAAAAGCAGCGGAG 690
 Qy 419 LeuGluArgGlnArgGluGluArgGlyGlnLeuGluArgGluAlaAlaLys 438
 Db 689 CTAGAACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 630

Qy 439 ArgGluLeuGluArgGlnArgGlnLeuGluTrpGluArgAsnArgArgGlnGlnLeuLeu 458
 Db 629 CGGGAACCTTGAAGGCAACGACAACTTGGTGGACCGAATCGAAGCGCAAGAACTACTA 570
 Qy 459 AsnGlnArgAsnLysGluGlnGlnAspIleValValLeuLysAlaLysLysThrLeu 478
 Db 569 AATCAAGAACAACAAAGAACAGAGGACATAGTTGTACTGAAGCAAAAGAAAGAACTTTG 510
 Qy 479 GluPheGluLeuGluAlaLeuAsnAspLysLysHisGlnLeuGluGlyLysLeuGlnAsp 498
 Db 509 GAATTTGAATTTAGAAGCTCTAAATGATPAAAGCATCACTAGAGGGAACCTTCAAGAT 450
 Qy 499 IleArgCysArgLeuThrThrGlnArgGlnGluIleGluSerThrAsnLysSerArgGlu 518
 Db 449 ATCAGATGTGATTTGACCCCAAGCAAGAAATTCAGAGCACAAACAAATCTAGAGAG 390
 Qy 519 LeuArgIleAlaGluIleThrHisLeuGlnGlnGlnLeuGlnSerGlnGlnMetLeu 538
 Db 389 TTGAGAATTTGCCGAATTCACCCATCTACAGCAACAATTACAGGAATCTCAGCAATGCTT 330
 Qy 539 GlyArgLeuIleProGluLysGlnIleLeuAsnAspGlnLeuLysGlnValGlnGlnAsn 558
 Db 329 GGAAGACTTATTCAGAAAACAGATACTCAATGACCAATTAAACAAAGTTCACAGAAC 270
 Qy 559 SerLeuHisArgAspSerLeuValThrLeuLysArgAlaLeuGluAlaLysGluLeuAla 578
 Db 269 AGTTTGACACAGAGATTCACTTGTGTACACTTAAAGAGCCTTAGAAGCAAAAGAACTAGCT 210
 Qy 579 ArgGlnHisLeuArgAspGlnLeuAspGluValGluLysGluThrArgSerLysLeuGln 598
 Db 209 CGGAGCACCCTACAGACCACCACTGGATGGAAGTGGAGAAAGAACTAGATCAAACTACAG 150
 Qy 599 GluIleAspIlePheAsnAsnGlnLeuLysGluLeuArgGluIleHisAsnLysGlnGln 618
 Db 149 GAGATTGATATTTCAATATATCAGCTGAGGAACCTAGAGAAATACACAATAAGCAACA 90
 Qy 619 LeuGlnLysGlnLysSerMetGluAlaGluArgLeuLysGlnLysGluGlnGluArgLys 638
 Db 89 CTCCAGAAGCAAAAAGTCCATGGAGGCTGAACGACTGAAACAGAAAGAAACAAAGAAAG 30
 Qy 639 IleIleGluLeuGluLysGlnLysGlu 647
 Db 29 ATCATGAATTAGAAAAAAGAAAAA 3

Search completed: August 3, 2004, 08:51:26
 Job time : 5740 secs